COGNITIVE PSYCHOLOGY **36,** 203–272 (1998)
ARTICLE NO. CG980687

Basing Categorization on Individuals and Events

Lawrence W. Barsalou

Emory University

Janellen Huttenlocher

University of Chicago

and

Koen Lamberts

University of Birmingham

Exemplar, prototype, and connectionist models typically assume that events constitute the basic unit of learning and representation in categorization. In these models, each learning event updates a statistical representation of a category independently of other learning events. An implication is that events involving the same individual affect learning independently and are not integrated into a single structure that represents the individual in an internal model of the world. A series of experiments demonstrates that human subjects track individuals across events, establish representations of them, and use these representations in categorization. These findings are consistent with ''representationalism,'' the view that an internal model of the world constitutes a physical level of representation in the brain, and that the brain does not simply capture the statistical properties of events in an undifferentiated dynamical system. Although categorization is an inherently statistical process that produces generalization, pattern completion, frequency effects, and adaptive learning, it is also an inherently representational process that establishes an internal model of the world. As a result, representational structures evolve in memory to track the histories of individuals, accumulate information about them, and simulate them in events.

We are grateful to Sarah Bunton, Catherine Norris, and Barbara Luka for laboratory assistance, and to Brian Ross, Judith Avrahami, Boris Brodsky, Yaakov Kareev, Jean Mandler, Douglas Medin, Gregory Murphy, Robert Nosofsky, Karen Solomon, and two anonymous reviewers for helpful comments on earlier drafts. This work was supported by National Science Foundation Grants SBR-9421326 and SBR 9796200 to Lawrence W. Barsalou.

Address correspondence and reprint requests to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, GA 30322. E-mail: barsalou@emory.edu.

0010-0285/98 $25.00

Current theories typically assume that events constitute the basic unit of categorization. As people process the members of a category, they store a memory of each categorization event. When assessing whether a subsequent entity belongs to the category, people retrieve these memories, assess their similarity to the new entity, and admit the entity to the category if sufficiently similar. Notably, this event-based approach ignores individuals. It does not assume that the cognitive system tracks the same individual across multiple events. Nor does it assume that the cognitive system stores information about individuals, or use such information in categorization decisions. Instead, this approach only assumes that information about events underlies categorization.

Figure 1 illustrates the centrality of events in current theories. Panel (a) presents an idealized representation of nine learning events. In each event, a member from the same category is encountered, and four features are extracted. For example, a learner might encounter a particular species of dogs on nine occasions, extracting four features on each. Alternatively, this idealized sequence might represent nine events with a particular tool category, or with a particular event category. An important property of the sequence in Fig. 1a is that the information encoded for five of the nine events is identical. For every odd-numbered event, the learner extracts the features $f_1f_2f_3f_4$. In contrast, for every even-numbered event, the learner extracts a unique feature set.

As Fig. 1b illustrates, this learning sequence is ambiguous. According to the *nonrepeating interpretation*, a unique individual occurs in each learning event—the same individual never repeats across events. Thus, the same feature set, $f_1f_2f_3f_4$, is extracted from five different individuals, I_1, I_2, I_3, I_4, and I_5. In contrast, according to the repeating interpretation, the same individual, I_1, repeats across these five events, with the same features extracted each time. In the remainder of this example, we focus on the repeating interpretation, because it allows us to contrast events-based and individuals-based approaches to categorization.

Three theoretical approaches currently dominate categorization research: exemplar models, prototype models, and connectionist models. Although all three typically adopt events as their basic unit of categorization, we will also consider versions of these models that adopt individuals as their basic units (these latter models are not widely entertained, but they are certainly feasible). For the sake of clarity, we only consider simple versions of these models that learn one category. However, our arguments readily generalize to more complex versions that learn multiple categories.

Exemplar Models

First consider the two exemplar models in Figs. 1d and 1e. Whereas the events model in Fig. 1d stores one exemplar for each learning event, the individuals model in Fig. 1e stores one exemplar for each individual, no matter how often it repeats. Nearly all exemplar models take the form in

(a) Training Sequence

E1	E2	E3	E4	E5	E6	E7	E8	E9	
f_1	f_1	f_1	f_2	f_1	f_3	f_1	f_4	f_1	
f_2	f_5	f_2	f_5	f_2	f_5	f_2	f_5	f_2	
f_3	f_6	f_3	f_6	f_3	f_6	f_3	f_6	f_3	
f_4	f_7	f_4	f_7	f_4	f_7	f_4	f_7	f_4	
I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	(NR)
I_1	I_2	I_1	I_3	I_1	I_4	I_1	I_5	I_1	(R)

(b) Non-repeating (NR) and Repeating (R) Interpretations of Individuals (above)

(c) Transfer Exemplars

T_1	T_2
f_1	f_5
f_2	f_6
f_8	f_8
f_9	f_9

(d) Events-Based Exemplar Memory

f_1	f_1	f_1	f_2	f_1	f_3	f_1	f_4	f_1
f_2	f_5	f_2	f_5	f_2	f_5	f_2	f_5	f_2
f_3	f_6	f_3	f_6	f_3	f_6	f_3	f_6	f_3
f_4	f_7	f_4	f_7	f_4	f_7	f_4	f_7	f_4

(e) Individuals-Based Exemplar Memory

f_1	f_1	f_2	f_3	f_4
f_2	f_5	f_5	f_5	f_5
f_3	f_6	f_6	f_6	f_6
f_4	f_7	f_7	f_7	f_7

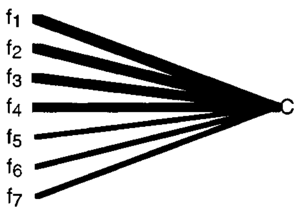
(f) Events-Based Prototype

f_1 - 6
f_2 - 6
f_3 - 6
f_4 - 6
f_5 - 4
f_6 - 4
f_7 - 4

(g) Individuals-Based Prototype

f_1 - 2
f_2 - 2
f_3 - 2
f_4 - 2
f_5 - 4
f_6 - 4
f_7 - 4

(h) Events-Based Connectionist Net



(i) Individuals-Based Connectionist Net

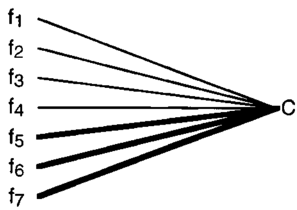


FIG. 1. Illustration of events-based and individuals-based models of categorization. (a) An idealized learning sequence of nine events, E_i. (b) Two transfer exemplars, T_i. (c) The nonrepeating(NR) interpretation of the learning sequence as containing nine individuals, I_i, and a repeating (R) interpretation of the learning sequence as containing five individuals. (d) Memory in an events-based exemplar model. (e) Memory in an individuals-based exemplar model. (f) An events-based prototype. (g) An individuals-based prototype. (h) An events-based connectionist net. (i) An individuals-based connectionist net.

Fig. 1d, assuming that each repetition of the same individual produces an independent trace in memory (e.g., Brooks, 1978; Heit, 1994; Hintzman, 1986; Lamberts, 1994; Medin & Schaffer, 1978; Nosofsky, 1984). Several notable exceptions that consider individuals include Medin and Edelson (1984), Nosofsky (1988, 1991), and Bareiss and Slator (1993).

To see the implications of storing events versus individuals, consider the two transfer exemplars in Fig. 1c. The events model prefers transfer exemplar $f_1f_2f_8f_9$ over transfer exemplar $f_5f_6f_8f_9$ because $f_1f_2f_8f_9$ overlaps more with the nine event memories in Fig. 1d. Whereas $f_1f_2f_8f_9$ shares two features with five memories and one feature with two memories, $f_5f_6f_8f_9$ only shares two features with four memories. In contrast, the individuals model prefers $f_5f_6f_8f_9$ over $f_1f_2f_8f_9$ because $f_5f_6f_8f_9$ overlaps more with the five individuals in Fig. 1e. Whereas $f_5f_6f_8f_9$ shares two features with four individuals, $f_1f_2f_8f_9$ only shares two features with one individual and one feature with two individuals.

As this example illustrates, storing events versus individuals has major implications. When events constitute the basic unit, the most frequent features across events control categorization. When individuals constitute the basic unit, the most frequent features across individuals are critical. To the extent that individuals differ in frequency of occurrence, the typical features for events models are not necessarily the same as the typical features for individuals models.

Prototype Models

Prototype models behave similarly to exemplar models. Figures 1f and 1g present prototype models based on events and individuals, respectively. In the events model, each repetition of the same individual increases the frequency of its features in the category prototype (Fig. 1f). As a result, the category prototype is $f_1f_2f_3f_4$. In contrast, the repeated individual's features only count once in the individuals model (Fig. 1g). As a result, the category prototype is $f_5f_6f_7$.

The different prototypes for the events and individuals models produce different transfer performance. Again consider the two transfer exemplars in Fig. 1c. The events model prefers $f_1f_2f_8f_9$ because it matches the events-based prototype, $f_1f_2f_3f_4$, on two features, whereas $f_5f_6f_8f_9$ matches it on none. In contrast, the individuals model prefers $f_5f_6f_8f_9$ because it matches the individuals-based prototype, $f_5f_6f_7$, on two features, whereas $f_1f_2f_8f_9$ matches it on none. Analogous to exemplar models, storing information about events versus individuals has major implications for categorization. Although prototype models based on individuals are feasible, prototype models are generally based on events (e.g., Reed, 1972; Smith & Medin, 1981).

Connectionist Models

Most connectionist models are also based on events (e.g., McClelland, Rumelhart, & the PDP Research Group, 1986; Rumelhart, McClelland, & the

PDP Research Group, 1986), although Kruschke (1992) provides a notable exception. Figures 1h and 1i present connectionist models based on events and individuals, respectively. In the events model, each repetition of the same individual increases the strength of the connections between the relevant feature units and the category unit (Fig. 1h). As a result, f_1, f_2, f_3, and f_4 develop the strongest connections to the category unit. In contrast, the repeated individual's features only strengthen these same connections once in the individuals model (Fig. 1i). As a result, f_5, f_6, and f_7 develop the strongest connections to the category unit.

The different connections in events models and individuals models produce different transfer performance. Essentially, each set establishes a different attractor in the network, leading to different preferences for the two transfer exemplars in Fig. 1c. The events model prefers $f_1f_2f_8f_9$ because it activates the attractor based on events, $f_1f_2f_3f_4$, better than does $f_5f_6f_8f_9$. In contrast, the individuals model prefers $f_5f_6f_8f_9$ because it activates the attractor based on individuals, $f_5f_6f_7$, better than does $f_1f_2f_8f_9$. Analogous to exemplar and prototype models, basing connection strengths on events versus individuals has major implications for categorization.

Optimizing Events versus Individuals during Categorization

When a model computes a category's typical features across events, it predicts the features most likely to characterize the category in a future *event.* It does not predict the features most likely to characterize a future *individual* of the category. To see this, consider a prediction of the events-based exemplar model in Fig. 1d. This model correctly predicts that feature f_1 will occur with a probability of .67 in future events that contain category members, given f_1 occurred in 6 of 9 previous events. In contrast, the individuals-based model in Fig. 1e does not accurately predict the probability of f_1 in a future event. Instead, its estimate of .40 reflects the proportion of individuals exhibiting f_1 regardless of how often they occur. For a model to produce optimal predictions of features in future events, it must consider all previous events, including repetitions of familiar individuals.

Conversely, when a model computes a category's typical features across individuals, it predicts the features most likely to characterize a future individual of the category. It does not predict the features most likely to characterize the category in a future event. Again consider the predictions of the models in Figs. 1d and 1e for f_1. If we seek an accurate prediction of f_1 in a future individual, the individuals-based model in Fig. 1e optimizes performance. Because this model computes the probability of a feature across individuals, it establishes the optimal prediction that f_1 will occur with a probability of .40. In contrast, the events-based model is nonoptimal, because it's inaccurate estimate of .67 is biased by how often particular individuals occur. For a model to produce optimal predictions of individuals, it must count each individual once and only once.

Neither approach to optimization is inherently more rational than the other. On some occasions, it may be useful to predict what will happen in the current event. On other occasions, it may be important to predict what will be true of the current individual. Although neither approach is superior to the other in principle, the human cognitive system may be predisposed to optimize one more than the other. The experiments to follow explore this issue.

Tracking and Representing Individuals

Before turning to the experiments, it is necessary address an important computational difference between these different approaches to categorization. Computationally, events models are simpler than individuals models. Regardless of whether we consider exemplar, prototype, or connectionist models, learning in events models is relatively simple. As Fig. 2a illustrates, events models simply encode the features of the current individual and then update category knowledge. As we saw in Fig. 1, the updating process can take the form of adding exemplars, revising prototypes, or adjusting connection weights. Most importantly, however, all three learning mechanisms do not vary their learning procedure as a function of whether the individual being processed is familiar (repeated) or unfamiliar (nonrepeated). Learning is not complicated by attempting to track individuals, nor by updating knowledge differently for familiar and unfamiliar individuals.

In contrast, individuals models require more complicated computational mechanisms. Regardless of whether we consider exemplar, prototype, or connectionist models, learning in individuals models requires tracking repeated individuals across events and establishing separate knowledge for each one. Figure 2b illustrates the additional mechanisms necessary for processing individuals. First, it is necessary to establish whether the current individual is familiar or unfamiliar. As Fig. 2b illustrates, two factors can affect this decision. First, a priori factors may predispose a perceiver toward believing that an about-to-be-seen individual is familiar or unfamiliar, as when expecting to see one's daily exercise partner at a prearranged meeting place. Even if the person doesn't look familiar or can't be seen clearly, the perceiver may infer, at least initially, that the individual is her exercise partner. Second, empirical factors are important as well. When not expecting to see anyone while hiking in the woods, a perceiver may nevertheless recognize an individual as familiar, based on physical features such as face and clothing.

After identifying an individual as familiar or unfamiliar, information about it is stored in memory. As Fig. 2b illustrates, the learning process differs for familiar and unfamiliar individuals, unlike events models. If the individual is familiar, no new memory representation is created. Instead, the existing representation for the individual is strengthened and updated as it guides processing of the individual in a top-down manner. In contrast, if the individ-

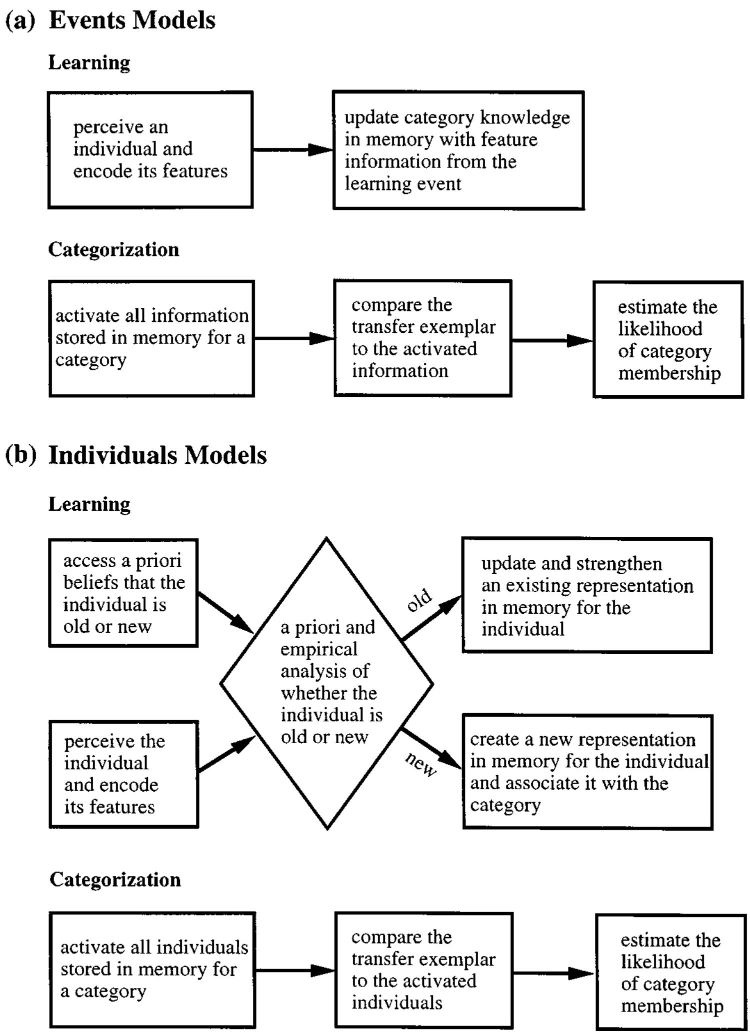


FIG. 2. Learning and categorization assumptions of (a) events models and (b) individuals models.

ual is unfamiliar, a new representation is created to represent the individual on this occasion and to accrue information acquired for it on future occasions. New representations are probably only created for individuals that are significant and processed deeply, not for individuals perceived briefly, such as cars passing by on the freeway.

The proposal that the cognitive system tracks and represents individuals is hardly novel. Indeed, it is a central assumption of traditional frame theory (e.g., Barsalou, 1992, 1993; Bartlett, 1932; Minsky, 1977; Rumelhart & Ortony, 1978) and of mental model theory (e.g., Johnson-Laird, 1983). To make this assumption explicit, Barsalou, Yeh, Luka, Olseth, Mix, and Wu (1993) dubbed it *the one-entity one-frame principle.* According to this principle, information extracted from an individual across multiple episodes is integrated into a single frame, assuming that the individual's identity is established correctly on each occasion. Across episodes, generic knowledge about an individual develops to organize episodic information into a single representational structure (Barsalou, 1988, 1992, 1993, in press). Similar accounts of category learning have been proposed by Thorndyke and Hayes-Roth (1979), Medin and Ross (1989), Ross, Perkins, and Tenpenny (1990), Spalding and Ross (1994), and Millikan (1998). Work on the type-token distinction is also relevant (e.g., Armstrong, 1989; Norman, Rumelhart, & the LNR Research Group, 1975).[1]

Several lines of evidence suggest that people establish integrated representations of individuals. In vision, repetitions of the same individual are tracked across perceptual events and integrated into a single object file (Treisman, 1992). A parallel integration process occurs at more cognitive levels. In social cognition, many researchers have shown that people integrate information about an individual across occasions into a single representation (e.g., Andersen & Cole, 1990; Andersen & Glassman, 1996; Bieke & Sherman, 1994; Rothbart, Fulero, Jensen, Howard, & Birrell, 1977; Srull & Wyer, 1989). For example, Srull and Brand (1983) presented subjects with randomly mixed information about two individuals and asked them to form an impression of each. Later, subjects' ability to remember descriptions of one individual exhibited within-individual interference but not between-individual interference, suggesting that subjects had constructed an integrated representation for each individual. Similarly, Thorndyke and Hayes-Roth (1979) and Watkins and Kerkar (1985) presented subjects with information about individuals across multiple occasions and found that subjects exhibited within-individual strengthening for common information and within-individual interference for unique information. Finally, recent developmental studies demonstrate that children track the same individual across repetitions (Bloom & Kelemen, 1995; Xu & Carey, 1996). Developmental studies that demonstrate the accumulation and integration of category information are closely related (e.g., Adler, 1997; Bauer & Fivush, 1992; Farrar & Goodman, 1992; Rovee-Collier, 1995).

Overview

Thus far we have explored two general classes of categorization models: Events models and individuals models. Experiments 1 and 2 evaluate a priori predictions of these two classes. When people learn a new category, do they

[1] Following Barsalou (1992), we assume that ''frame'' and ''schema'' are different names for the same construct. We use ''frame'' here to maintain continuity with previous work.

behave like an events model or an individuals model? As we will see, the results are mixed. Some results clearly indicate that people store events and attempt to optimize them during categorization. Other results, however, suggest that people track individuals. To accomodate this complex pattern, we develop the class of *individuals sampling models.* Although these models track and represent individuals, they also store and optimize information about events. In Experiments 3 and 4, we test a priori predictions of these models. Finally, the General Discussion explores other hybrid models, as well as issues related to individuals and events in categorization.

The Basic Paradigm

Previous research has found that frequent stimuli have more impact on categorization than infrequent stimuli (e.g., Barsalou, 1981, 1985; Florian, 1992; Heit & Barsalou, 1996; Huttenlocher, Hedges, Engebretson, & Vevea, 1996; Nosofsky, 1988, 1991; Rips & Collins, 1993). Although this research might seem to support events models, it contains an ambiguity that allows individuals models to explain the observed frequency effects as well. Subjects in these studies were free to perceive the multiple presentations of a stimulus as either different individuals or as repetitions of the same individual. If a repeating stimulus was perceived as different individuals, then individuals models predict frequency effects, because they assume that more individuals are represented in memory for a frequent stimulus than for an infrequent one. Under this interpretation, individuals models make the same predictions as events models.

To distinguish these classes of models, we explicitly manipulated whether subjects believed a repeating stimulus was a single individual or different individuals. Of particular interest was the categorization performance of subjects who believed that each repeating stimulus was a single individual. When the presentation frequency of these individuals differed, did subjects nevertheless weight them equally during categorization, as individuals models predict? Or did subjects allow event frequency to affect categorization, as events models predict?

In the experiments to follow, subjects studied five exemplars from a fictional category of tropical fish. Figure 3 presents the concrete exemplars that subjects studied. Table 1 presents the abstract structure that underlies these exemplars. As Fig. 3 and Table 1 illustrate, Exemplars 1, 2, 3, and 4 were similar to one another, whereas Exemplar 5 was different. As Table 1 further illustrates, these exemplars varied in presentation frequency, with Exemplars 1, 2, 3, and 4 occurring 3 times each, and Exemplar 5 occurring 18 times.[2]

[2] We use *exemplar* as a methodological term that refers only to stimulus items in Tables 1 and 4. We do not use *exemplar* as a theoretical term that refers to cognitive representations. Instead, *event memory* and *individual* serve this function, referring to the two possible ways in which people could represent a stimulus item.

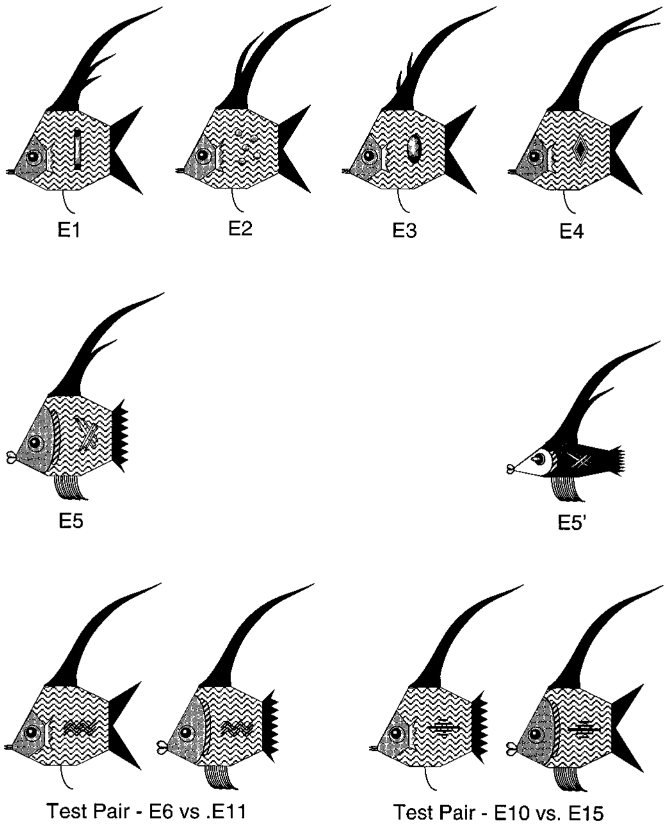


FIG. 3. Top row: The four similar individuals presented during learning (E1, E2, E3, and E4). Middle row: The frequent dissimilar exemplar presented during learning in every condition of every experiment (E5), except in the far-similarity condition of Experiment 3, where E5′ was the frequent dissimilar exemplar. Bottom row: Examples of two test pairs.

To assess subjects' use of individuals in categorization, each experiment manipulated whether subjects regarded individuals as repeating or nonrepeating. In the *nonrepeating condition,* subjects were led to believe that each of the 30 presentations reflected a different individual from the category. Because 60% of these presentations took the form of Exemplar 5, subjects should have believed that a fish similar to Exemplar 5 was more likely to be a category member than a fish similar to Exemplars 1 through 4, which constituted only 40% of the presentations. Because fish similar to Exemplar 5 constituted 60% of the events *and* 60% of the individuals, both events models and individuals models make this prediction. On both accounts, nonrepeating subjects should have preferred transfer Exemplars 11, 12, 13, 14, and 15 in Table 1 when paired with transfer Exemplars 6, 7, 8, 9, and 10, respectively (Fig. 3 provides two examples of these test pairs).

TABLE 1
Category Structure in Experiments 1, 2, 3, and 4 (Except for the Far-Similarity Condition of Experiment 3)

Exemplar	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Presentation frequency
	Acquisition exemplars										
1	1	1	1	1	1	1	1	1	1	1	3
2	2	2	1	1	1	1	1	1	1	1	3
3	3	3	1	1	1	1	1	1	1	1	3
4	4	4	1	1	1	1	1	1	1	1	3
5	5	5	0	0	0	0	1	1	1	1	18 (3)[a]
Exemplar	D1	D2[b]	D3	D4	D5	D6	D7	D8	D9	D10	Contrast exemplar
	Transfer exemplars										
6	6	6/7	1	1	1	1	1	1	1	1	11
7	6	6/7	1	1	1	0	1	1	1	1	12
8	6	6/7	1	1	0	1	1	1	1	1	13
9	6	6/7	1	0	1	1	1	1	1	1	14
10	6	6/7	0	1	1	1	1	1	1	1	15
11	6	6/7	0	0	0	0	1	1	1	1	6
12	6	6/7	0	0	0	1	1	1	1	1	7
13	6	6/7	0	0	1	0	1	1	1	1	8
14	6	6/7	0	1	0	0	1	1	1	1	9
15	6	6/7	1	0	0	0	1	1	1	1	10

[a] Exemplar 5 had a presentation frequency of 18 in every condition of every experiment, except for the equal frequency conditions of Experiment 1, where it had a frequency of 3.
[b] The two exemplars in each transfer pair always had the same value on D2, either 6 or 7.

Events models and individuals models make contrasting predictions for the *repeating condition.* These subjects were led to believe that the 30 presentations contained multiple repetitions of the same individuals, with some individuals more likely to occur than others. Under these instructions, subjects should have believed that five individual fish occurred multiple times, with Exemplars 1 through 4 occurring occasionally and Exemplar 5 occurring frequently. If repeating subjects based their categorizations on equally weighted individuals, as individuals models predict, they should have believed that a new fish similar to Exemplars 1, 2, 3, and 4 was more likely to be a category member than a fish similar to Exemplar 5. When individuals are the unit of analysis, Exemplars 1, 2, 3, and 4 outnumber Exemplar 5. In contrast, if repeating subjects based their categorizations on events, as events models predict, they should have believed that a fish similar to Exemplar 5 was more likely to be a category member than a fish similar to Exemplars 1, 2, 3, and 4. When events are the unit of analysis, the 18 events for Exemplar 5 outnumber the 12 total events for Exemplars 1, 2, 3, and 4.

Besides using transfer preferences to assess events and individuals models, we also used two other types of data. First, we examined the performance of individual subjects. Do all subjects uniformly optimize events or individuals, or do subjects differ, with some optimizing events and others optimizing individuals? Second, we examined subjects' memory for the frequency of events and individuals. Do subjects acrue frequency information for events, individuals, or both? If the cognitive system develops to primarily perform one type of optimization, subjects should adopt it uniformly, and they should only store frequency information relevant to it. On the other hand, if optimization is flexible, or if it operates simultaneously on both events and individuals, subjects may adopt multiple strategies, and they may exhibit frequency sensitivity to both types of information.

EXPERIMENT 1

The first experiment provided an initial assessment of whether subjects base categorization on events or individuals. Two factors varied orthogonally between subjects. First, as just described, repeating subjects believed that 5 individuals occurred multiple times during learning, whereas non-repeating subjects believed that 30 individuals occurred once. Second, the frequency of the five individuals was unequal, as just described, or equal. When frequency was unequal, Exemplars 1, 2, 3, and 4 each occurred 3 times, and Exemplar 5 occurred 18 times. When frequency was equal, all 5 exemplars occurred 3 times each. Crossing these two factors produced four conditions: repeating / unequal, repeating / equal, nonrepeating / unequal, and nonrepeating / equal.

The repeating / unequal condition was of primary interest. Events models predict that Exemplar 5 should dominate categorization, because it occurs in the most events. In contrast, individuals models predict that the Exemplars 1, 2, 3, and 4 should dominate, because their four frames outnumber the one frame for Exemplar 5.[3]

The remaining three conditions evaluate two assumptions of the basic paradigm. The first assumption is that the frequency manipulation is sufficiently potent to produce the frequency dominance predicted by events models. The nonrepeating / unequal condition assesses this assumption. If the frequency manipulation is potent, then both models predict that Exemplar 5 should dominate categorization. According to both models, subjects should allow the 18 individuals that take the form of Exemplar 5 to dominate the 12 total individuals that take the form of Exemplars 1, 2, 3, and 4.

The second assumption is that the confounding of similarity with frequency in the design is not responsible for any observed frequency domi-

[3] Later, we will show that the intuitive predictions in this paragraph for events and individuals models hold robustly across a variety of formal categorization models.

nance. To see this, imagine that Exemplar 5 dominates categorization in the repeating / unequal condition, as events models predict. Exemplar 5 could dominate, either because it occurs most frequently, or because there is something about its dissimilar features that makes it more salient or desirable. The two equal-frequency conditions assess this issue by eliminating the frequency differential between Exemplar 5 and the other exemplars. If Exemplar 5's features are responsible for what appears to be frequency dominance, then Exemplar 5 should continue to dominate. If subjects prefer Exemplar 5 because it is salient or desirable, they should still prefer it when it doesn't have a frequency advantage. Alternatively, if Exemplars 1, 2, 3, and 4 dominate Exemplar 5 in the equal-frequency conditions, then Exemplar 5's dominance in the unequal-frequency conditions reflects its high frequency and not other factors.[4]

Method

Design and subjects. Two between-subject variables and one within-subject variable were crossed to structure the experiment. The between-subject variables were repeating vs nonrepeating training instructions, and equal vs unequal exemplar frequency. The within-subject variable was the manipulation of 1111 versus 0000 transfer exemplars, as described in the materials section. The dependent measures were categorization choices, frequency estimates, and typicality rankings, as described shortly. Subjects were 48 members of the University of Chicago community, who received $2.00 for 15 to 20 min of participation. Twelve subjects were assigned randomly to each of the four training instruction × exemplar frequency cells of the design. The different versions of the training and test materials described next were distributed equally across the experimental variables.

Materials. During the training phase, subjects studied individuals from a fictional species of tropical fish. The individuals studied were repetitions of Exemplars 1, 2, 3, 4, and 5 in Fig. 3. As Table 1 in conjunction with Fig. 3 illustrates, 10 dimensions structured the exemplars: top fin (D1), side marking (D2), tail (D3), bottom fin (D4), face shape (D5), mouth (D6), eyes (D7), face color (D8), body pattern (D9), and body shape (D10). Dimensions 1 and 2 individuated the training exemplars, with each exemplar having a unique value on each dimension. Dimensions 3, 4, 5, and 6 created a similarity gradient between Exemplars 1, 2, 3, and 4 and Exemplar 5. Whereas Exemplars 1, 2, 3, and 4 possessed the 1 value for each of these dimensions, Exemplar 5 possessed the 0 value. Dimensions 7, 8, 9, and 10 were common to all training and transfer exemplars. These dimensions become relevant in Experiment 3 when we address the issue of Exemplar 5's similarity to Exemplars 1, 2, 3, and 4. As Table 1 further indicates, the unequal frequency materials contained 18 repetitions of Exemplar 5 and 3 repetitions of Exemplars 1, 2, 3, and 4, whereas the equal frequency materials contained 3 repetitions of each exemplar. The 30 or 15 training exemplars in a given set of materials appeared in one of two random orders.

Table 1 also describes the ten transfer exemplars. We will call Exemplars 6, 7, 8, 9, and 10 the *1111 transfer exemplars,* because they have 3 or 4 values of 1 on dimensions 3, 4, 5, and 6. We will call Exemplars 11, 12, 13, 14, and 15 the *0000 transfer exemplars,* because they have 3 or 4 values of 0 on these dimensions. The 1111 transfer exemplars include the

[4] In designing the equal-frequency conditions, we could have presented each exemplar 6 times and held the total number of presentations constant at 30. Because we wanted to ensure that 3 presentations were sufficient for learning Exemplars 1, 2, 3, and 4, however, we held the number of presentations constant at 3 and reduced the total number of presentations to 15.

1111 ‘‘prototype’’ (Exemplar 6) and four exemplars that differ systematically on one dimension from it (Exemplars 7, 8, 9, and 10). The 0000 transfer exemplars include the 0000 ‘‘prototype’’ (Exemplar 11) and four exemplars that differ systematically on one dimension from it (Exemplars 12, 13, 14, and 15). As can be seen, the transfer exemplars all had the same top fin (value 6 for D1), which differed from the top fins for all of the training exemplars. Additionally, the transfer exemplars also had one of two new side markings (value 6 or 7 on D2), which also differed from all of the training exemplars.

On the choice test, subjects received pairs of transfer exemplars in a forced-choice format, each pitting a 1111 transfer exemplar against a matched 0000 transfer exemplar. Table 1 lists the contrast exemplar for each transfer exemplar, and Fig. 3 presents two of these test pairs. As Table 1 illustrates, the 10 test pairs contain two replications of five critical pairs: Exemplar 6 vs Exemplar 11, Exemplar 7 vs Exemplar 12, Exemplar 8 vs Exemplar 13, Exemplar 9 vs Exemplar 14, and Exemplar 10 vs Exemplar 15. These five pairs are highly informative because they maximize the ability to distinguish events models and individuals models, and because they facilitate estimating dimension weights (Appendix B). Each pair occurred twice. The first replication provided an independent data point for one transfer exemplar in the pair, and the second replication provided an independent data point for the other transfer exemplar. For example, if a subject’s preference for Exemplar 6 was assessed in the first replication of the pair containing Exemplars 6 and 11, then the subject’s preference for Exemplar 11 was assessed independently in the second replication.

To decrease the likelihood that subjects would perceive the two replications of each critical pair as related, the replications differed in two ways: First, the two transfer exemplars were inversely ordered in the replications, with each exemplar appearing once on the left and once on the right. Second, the side markings differed for each replication. In one replication, the two transfer exemplars both had value 6 on D2 (side markings), and in the other they both had value 7 (in Fig. 3, the E6 vs E11 test pair illustrates value 6, and the E10 vs E15 test pair illustrates value 7). Thus, the reversals of position and side marking across the two replications obscured the fact that the critical exemplars were the same. Four versions of the 10 choice pairs were constructed such that each of the 10 transfer exemplars occurred equally often on the left or right, equally often with value 6 or 7 as its side marking, and equally often in the first or second replication of a critical pair. Within each version, the first half of the list contained the first replications of the five critical pairs in a random order, and the second half contained the second replications in a different random order, with the constraint that the fifth and sixth pairs not be replications of the same critical pair.

On the typicality ranking test, subjects received all ten transfer exemplars without side markings on a single page in one of two random orders.

Procedure. Prior to the training phase, subjects were asked to imagine that they were viewing a species of tropical fish at Chicago’s Shedd Aquarium. Nonrepeating subjects were shown a diagram (Fig. 4a) that illustrated a series of tanks: the source tank, the viewing aquarium, and the destination tank. Subjects were told that all fish began the day in the source tank and swam one by one through the viewing aquarium to the destination tank. The cover story explained that the fish could only be viewed one at a time, else they became agitated and injured each other. A series of interlocking one-way gates prevented multiple fish from entering the viewing area simultaneously and from returning after their single visit. The key point conveyed was that subjects would see each fish once and only once. Subjects were also told that some of the fish would look very similar to each other but that they would be different individuals.

Subjects in the repeating condition were shown a different diagram (Fig. 4b) that illustrated a holding tank and a viewing area. Subjects were again told that the fish became agitated when viewed and that they therefore could only be viewed one at a time. In this arrangement of tanks, however, a given fish could swim out of the holding tank and into the viewing area multiple times, with interlocking one-way gates preventing multiple fish from entering the viewing area simultaneously. Most importantly, subjects were told that the fish would enter

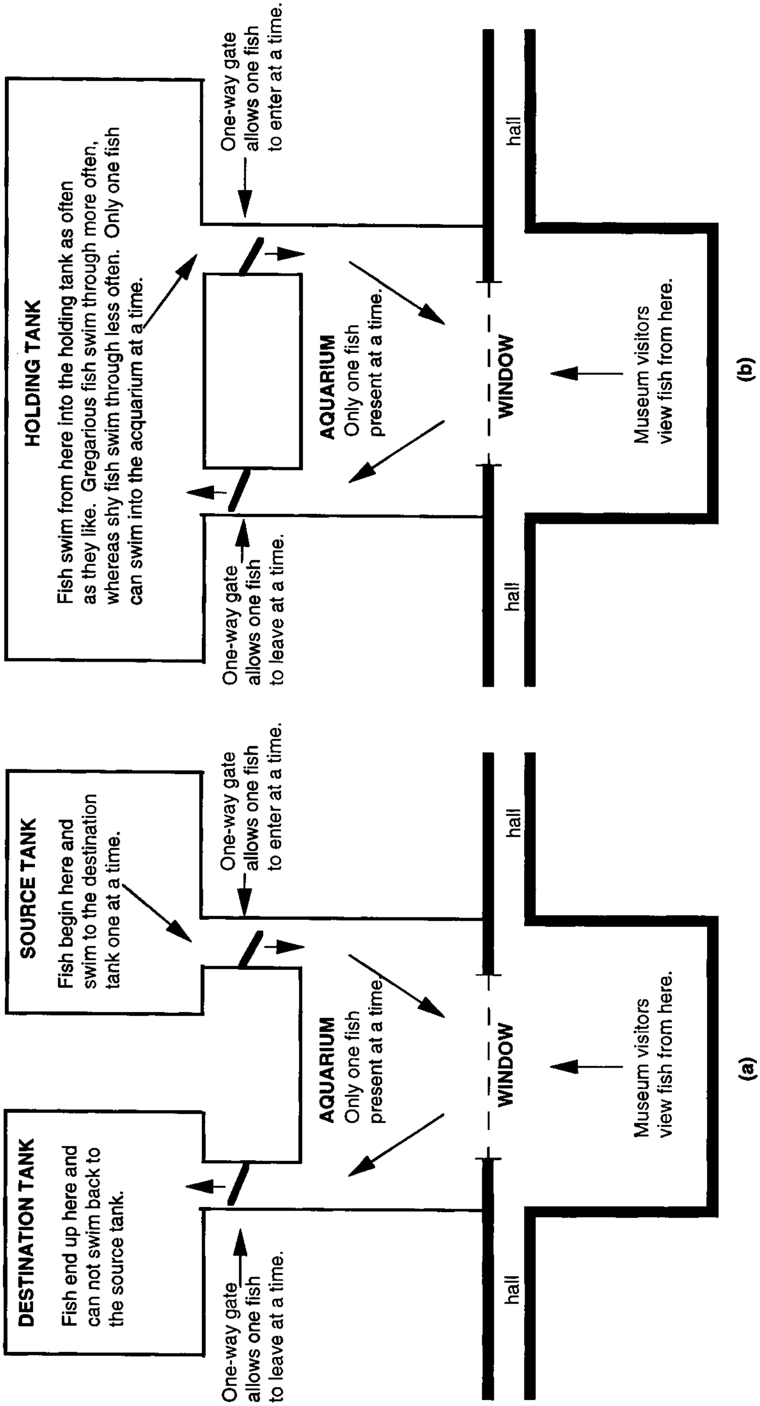


FIG. 4. Aquarium systems in the cover stories for the (a) nonrepeating and (b) repeating conditions.

the viewing area multiple times. Because the fish varied in how gregarious they were, however, some would swim through more often than others.

Following the cover story, subjects were told that their task was to learn as much as they could about the fish in the aquarium and that they would be asked some questions about these fish later. Subjects were told about the six critical dimensions that distinguished individual fish from one another, using a diagram that depicted these dimensions' positions on the fish but not illustrating their values. Subjects then observed 30 or 15 fish, each 10 cm tall, on a 11.5 by 8 cm laminated card, in a random order, at a 15 s presentation rate.

After training, subjects performed three transfer tasks in the following order: choice, frequency, and typicality. In the choice task, subjects received the ten randomly-ordered pairs of transfer exemplars and had to select which exemplar in a pair was more likely to be a category member. Each pair was shown on a separate 21.5 $\times$ 14 cm page, with each fish being 10 cm tall. At the top of each page appeared the query, ''Which of these two new fish is more likely to belong to the same species of fish that you saw earlier?'' To indicate their response, subjects placed an X below their choice. Next, subjects performed a frequency task, estimating ''the total number of times that a fish swam through the viewing area'' during training (event frequency). Subjects in the repeating condition also estimated the total number of individual fish that they had observed, independent of how often they had seen each one (individual frequency). Finally, in the typicality ranking task, subjects received the ten transfer exemplars, each 6 cm tall, randomly ordered on a 21.5 by 28 cm page. Subjects were asked to write 1 below the best example of the category, 10 below the worst example, 2 below the next best example, 9 below the next worst example, iterating in towards finally ranking the fifth and sixth best examples.

Results and Discussion

Choice proportions and typicality rankings. The critical results concern performance in the repeating / unequal condition. If individuals models are correct, subjects in the repeating / unequal condition should prefer the 1111 transfer exemplars over the 0000 transfer exemplars, selecting them more often on the choice task and judging them higher in typicality. Even though Exemplar 5 occurred more often than Exemplars 1, 2, 3, and 4, subjects should integrate all episodes for an individual into a single frame and ignore frequency, such that each individual has equal impact on subsequent categorization decisions. Because four individuals had 1 values on dimensions 3, 4, 5, and 6, they should dominate the one individual having 0 values on these dimensions. On the other hand, if events models are correct, subjects should prefer the 0000 transfer exemplars, exhibiting frequency dominance. Because there are more event memories for Exemplar 5 than for Exemplars 1, 2, 3, and 4 combined, subjects should prefer transfer exemplars similar to Exemplar 5.

Table 2 presents the relevant findings. Each 1111 choice proportion is the proportion of trials on which subjects chose the 1111 transfer exemplar in the five test pairs that provided data on the five 1111 transfer exemplars. Each 0000 choice proportion is the analogous proportion across the five test pairs that provided data on the five 0000 transfer exemplars. The 1111 and 0000 proportions in each condition need not sum to one, given they are, in principle, independent. Conceivably, subjects could have selected all 10 target exemplars across the 10 pairs, in which case both proportions would

TABLE 2
Choice Proportions, Typicality Rankings, Model Fits, and Frequency Estimates for Experiment 1

Measure/Factor	Unequal frequency: Repeating	Unequal frequency: Nonrepeating	Equal frequency: Repeating	Equal frequency: Nonrepeating
Choice proportions				
1111 Transfer exemplars	.22	.13	.80	.88
0000 Transfer exemplars	.75	.92	.13	.10
Typicality rankings				
1111 Transfer exemplars	6.33	5.90	4.95	5.03
0000 Transfer exemplars	4.67	5.07	6.05	5.97
Choice model fits				
Unweighted linear model	−.93	−.95	.98	.96
Weighted linear model	.93	.95	.98	.96
Typicality model fits				
Unweighted linear model	−.84	−.34	.53	.43
Weighted linear model	.84	.34	.53	.43
Frequency estimates				
Events	28.50	26.08	16.75	15.50
Individuals	7.83	—	7.17	—

have been one. When the two proportions do not sum exactly to one, it either indicates inconsistency in a subject's strategy, or a subject's insensitivity to the distinction between 1111 and 0000 transfer exemplars. Each entry for the typicality rankings in Table 2 is the average ranking across either the 1111 or 0000 transfer exemplars, where 1 was most typical and 10 was least typical. For readers interested in the average choice proportions and typicality rankings for individual transfer exemplars, Appendix A presents these data for all four experiments.

To assess the critical hypotheses, four summary scores were computed for each subject: the proportion of the five 1111 transfer exemplars selected, the proportion of the five 0000 transfer exemplars selected, the average typicality ranking of the 1111 transfer exemplars, and the average typicality ranking of the 0000 transfer exemplars. Two ANOVAs were then performed, one for choice and one for typicality. For choice proportions, the data were transformed using an arcsin transformation (Winer, 1971), with all MS_es for F tests reported in units on this transformed scale. The choice and typicality data in all later experiments were also analyzed in this manner.

As the findings in Table 2 illustrate, the results strongly favor events models. Most critically, subjects in the repeating / unequal condition preferred the 0000 transfer exemplars over the 1111 transfer exemplars ($F(1, 44) = 18.03$, $MS_e = .44$, $p < .01$). Although the 0000 transfer exemplars were generally ranked as more typical than the 1111 transfer exemplars, this difference was only marginally significant ($F(1, 44) = 3.64$ $MS_e = 4.57$, $p <$

.10). Together the choice and typicality data indicate that event frequency influences subjects' categorization judgments. Subjects preferred transfer exemplars similar to one frequent individual more than transfer exemplars similar to four similar but less frequent individuals.

Results from the other three conditions address various issues about the design. First, it is clear that the frequency manipulation was potent. In a combined analysis of the two unequal-frequency conditions, the more frequent 0000 transfer exemplars dominated the less frequent 1111 transfer exemplars, both for choice ($F(1, 44) = 40.89$, $MS_e = .44$, $p < .01$) and typicality ($F(1, 44) = 4.08$, $MS_e = 5.47$, $p < .05$). Second, it is clear that frequency—not a variable correlated with frequency—underlies the preference for 0000 transfer exemplars in the unequal-frequency conditions. When all five exemplars were presented equally often in the equal-frequency conditions, subjects no longer favored the 0000 transfer exemplars. If Exemplar 5 had been salient or desirable for some unanticipated reason, it should have continued to dominate transfer. Clearly, however, subjects in the equal-frequency conditions allowed Exemplars 1, 2, 3, and 4 to dominate. On the choice task, these subjects preferred 1111 transfer exemplars ($F(1, 44) = 54.57$, $MS_e = .44$, $p < .01$); on the typicality task, the advantage for 1111 transfer exemplars approached significance ($F(1, 44) = 2.78$, $MS_e = 4.57$, $p = .12$). When frequency was not a factor, subjects allowed the four similar individuals to dominate categorization.

Model-based analysis of the data. Intuitively, it may seem obvious that events models predict a 0000 preference in the repeating / unequal condition, whereas individuals models predict a 1111 preference. However, it is useful to verify these predictions formally (indeed, the authors disagreed on them initially). Because these models can incorporate a variety of nonlinear mechanisms, and because they can take many forms in their parameter spaces, it is possible that not all instances of events models conform to our intuitions. To explore this issue, we fit four models to the data in every experiment: Reed's (1972) linear exemplar model, Medin and Schaffer's (1978) context model, Nosofsky's (1984) generalized context model, and Lamberts' (1994) weighted ratio model. For each model, we developed two versions: In the *weighted* version, each presentation of an individual established an independent event memory, such that frequent individuals accrued more memories than infrequent individuals. In the *unweighted* version, only a single unweighted representation of an individual became established in memory, regardless of how often the individual occurred.[5]

[5] It is important to note that the weighted and unweighted versions of these four models are *not* equivalent to events and individuals models. Instead, the weighted and unweighted models are statistical tools that simply assess whether subjects weighted individuals by frequency or weighted them equally. In other words, the weighted and unweighted models only assess one assumption of more complex events and individuals models, namely, how they

Across every experiment, each of the four models confirmed our intuitive predictions: When subjects preferred the 0000 transfer exemplars, the weighted version of each model provided the best fit. In contrast, when subjects preferred the 1111 transfer exemplars, the unweighted versions were always superior. Furthermore, and perhaps surprisingly, all four models produced equally good fits of the data, with no one model dominating. We therefore report only fits of the linear exemplar model for two reasons. First, it's the only parameter-free model, with all other models having one or more free parameters.[6] Furthermore, in other work (Lamberts & Barsalou, 1998), we assess the fit of non-linear models, such that reporting the linear model here demonstrates the generality of our findings.

Appendix B presents the details of the linear exemplar model. To follow the modeling results here, however, it is only necessary to know that the unweighted and weighted versions of this model differed in their treatment of exemplar frequency. In the unweighted linear model, each exemplar was represented only once in memory, regardless of how often it occurred. Thus, in the repeating / unequal condition, this model established five representations for the five exemplars. In contrast, the weighted linear model represented each presentation of the same exemplar as an independent event memory. Thus, in the repeating / unequal condition, this model established 30 memories, 3 each for Exemplars 1, 2, 3, and 4, and 18 for Exemplar 5. Fitting the unweighted and the weighted linear models to the same data allowed us to determine whether repeating subjects used individuals or events as the unit of categorization. If subjects in the repeating / unequal condition used individuals, then the unweighted linear model should fit their data better than the weighted linear model. In contrast, if subjects used events, then the weighted linear model should fit better.

Because we were not interested in assessing response mechanisms, and because we wanted to minimize the number of free parameters, we did not attempt to reproduce subjects' exact responses. Instead, we only attempted to model the underlying similarity relations that subjects computed between transfer and training exemplars. Thus, the appropriate measure of fit was correlation, not squared deviations, and this is the measure that we report in all four experiments.

As Table 2 illustrates, fits of the weighted and unweighted linear models confirm the results for choice and typicality. In both unequal-frequency conditions, the weighted linear model provides the best fit, indicating that these subjects were sensitive to event frequency. Even when subjects knew that

---

weight individuals. As we shall see later, it's possible for certain individuals models to predict frequency effects, such that high fits of weighted models constitute support for them.

[6] For all models, including the linear exemplar model, dimension weights were *not* free parameters. As described in Appendix B, the structure of the transfer test allowed empirical assessments of dimension weights that were model independent.

individuals repeated, they weighted Exemplar 5 more heavily than Exemplars 1, 2, 3, and 4. In the equal-frequency conditions, both models fit equally well, because equating presentation frequency eliminated the basis for differential predictions. As already described, the unweighted linear model predicts a 1111 preference, because Exemplars 1, 2, 3, and 4 outnumbered Exemplar 5 as individuals. Similarly, the weighted linear model now predicts a 1111 preference, because the 12 total presentations of Exemplars 1, 2, 3, and 4 outnumber the 3 presentations of Exemplar 5.[7]

Analysis of individual subjects. Although the group data strongly support events models, the performance of individual subjects suggests caution in drawing a hasty conclusion. Of the 12 subjects in the repeating / unequal condition, 9 preferred 0000 transfer exemplars, and 3 preferred 1111 transfer exemplars.[8] Consistent with the group data, a majority of the subjects based categorization on events. Critically, however, if the cognitive system is inherently predisposed to optimize events during categorization, all 12 subjects should have preferred 0000 transfer exemplars. Because 3 subjects preferred 1111 transfer exemplars, the human cognitive system is clearly capable of basing categorization on individuals as well as on events. This finding provides the first hint that neither events nor individuals models are satisfactory, and that a more complicated hybrid model will be necessary.

Frequency estimates. As Table 2 shows, subjects were reasonably accurate at estimating the total number of training events. Subjects' estimates were near the correct value in both the unequal-frequency conditions (30) and in the equal-frequency conditions (15). Most critically, the frequency estimates for individuals provide the second indication that repeating subjects represented individuals. If they had not, their frequency estimates for individuals should have approximated their frequency estimates for events. In the unequal-frequency condition, repeating subjects should have believed that about 30 individuals occurred; in the equal-frequency condition, repeating subjects should have believed that about 15 individuals occurred. However, both groups estimated around 7 individuals, which is much closer to the actual number of individuals (5) than to the actual numbers of events (30 or 15). Furthermore, the unequal- and equal-frequency conditions did not differ in their estimates of individuals ($F(1, 22) < 1$, $MS_e = 13.26$), but did

[7] The unweighted and weighted linear models always produce symmetric fits. As Table 2 illustrates, the fits for both models have the same absolute value but are opposite in sign (except for the equal-frequency conditions, where the signs are the same). The last paragraph of Appendix B explains why the design of the experiments entails this symmetric pattern of fits.

[8] In all four experiments, each subject's choice preference was established by subtracting the number of 0000 choices for the five 0000 transfer pairs from the number of 1111 choices for the five 1111 transfer pairs. Subjects receiving a score of +2 or greater were coded as having a 1111 preference; subjects receiving a score of −2 or less were coded as having a 0000 preference; subjects receiving scores of 1, 0, or −1 were coded as having no preference.

differ in their estimates of events ($F(1, 22) = 62.46$, $MS_e = 13.26$, $p < .01$), with the interaction between frequency condition and frequency estimate being significant ($F(1, 22) = 27.79$, $MS_e = 13.26$, $p < .01$). This dissociation between frequency condition and frequency estimate suggests that subjects tracked individuals independently of events.

Some concern might arise because repeating subjects tended to overestimate the number of individuals by about two to three individuals. Perhaps these overestimates indicate that subjects did not track individuals properly across repetitions. If subjects did not establish representations of the proper individuals, then the assumptions for testing individuals models may not have been met, and the conclusions drawn from the choice and typicality data may be invalid. Experiment 2 addresses this issue by helping subjects track individuals.

EXPERIMENT 2

Perhaps most subjects didn't base categorization on individuals in Experiment 1 because it wasn't clear who the individuals were. When an individual repeated after its first presentation, subjects may have misconstrued it as a new individual or miscategorized it as the wrong individual, at least on some trials. As a result, subjects may not have established integrated representations in memory for the individuals, such that they could process them properly during categorization. This problem may have been particularly acute for the four similar individuals that shared many common features, although each of them did have two unique features that were salient (i.e., top fins, D1, and side markings, D2).

To address this issue, subjects in Experiment 2 received cues that helped them track repetitions of the same individual. When subjects can easily recognize the same individual across repetitions, do they use individuals during categorization rather than events? Subjects received the same exemplars as in Experiment 1, with Exemplars 1, 2, 3, and 4 occurring 3 times each and Exemplar 5 occurring 18 times. In the critical conditions, subjects received individuating cues that helped them track each individual's repetitions. Of interest was whether these cues eliminated frequency dominance.

The experiment included six between-subject conditions. Two conditions served as baselines for four critical conditions that contained individuating information, and they also provided an opportunity to replicate the two unequal-frequency conditions of Experiment 1. In the nonrepeating baseline, subjects received nonrepeating instructions and no individuating cues. In the repeating baseline, subjects received repeating instructions and no individuating cues. In the remaining four conditions, subjects received repeating instructions and individuating cues. In the number condition, subjects were told, during the acquisition instructions, that they would study five individuals who would repeat across trials. In the side markings condition, subjects were told that the side markings of the fish individuated them. In the names condition, subjects were told that each individual had a unique name, and that the names would be printed on the cards for each individual during acquisition. In each of these three conditions, only one of the three types of individuating information was included, excluding the other two. Of interest was the relative extent to which each type of information diminished frequency dominance, relative to the repeating / no-individuation baseline. Finally, subjects in the all-individuation condition received all three types of individuating information. Of interest was whether these subjects would exhibit any frequency dominance at all.

Method

Design and subjects. One between-subject variable and one within-subject variable were crossed to structure the experiment. The between-subject variable manipulated the acquisition instructions across six groups of subjects. As just described, these six groups were nonrepeating / no individuating cues, repeating / no individuating cues, repeating / number, repeating / side markings, repeating / names, and repeating / all individuating cues. The within-subject variable contrasted 1111 transfer exemplars with 0000 transfer exemplars, as in Experiment 1. The dependent measures were again categorization choices, frequency estimates, and typicality rankings. Subjects were 96 members of the University of Chicago community, who received $2.00 for 15 to 20 min of participation. Sixteen subjects were assigned randomly to each of the six instruction cells of the design. The different versions of the training and test materials described for Experiment 1 were distributed equally across the experimental variables.

Materials. Subjects received the same acquisition exemplars as in Experiment 1, with Exemplars 1, 2, 3, and 4 presented 3 times each, and Exemplar 5 presented 18 times. In the names and all-individuation condition, each training exemplar was assigned a name (Angela, Clarissa, Vertna, Charmaign, or Lois). On each presentation of the same training exemplar, the same name appeared in the upper left corner of the stimulus card. No names appeared for the transfer exemplars. All other aspects of the training and transfer materials were the same as in Experiment 1.

Procedure. In the conditions that provided individuating information, the relevant individuating information was stated saliently multiple times during the acquisition instructions. All other aspects of the procedure were the same as Experiment 1.

Results and Discussion

Choice proportions and typicality rankings. Table 3 presents the average choice proportions and typicality rankings from Experiment 2. The data from the two conditions that lack individuating information replicate the two unequal-frequency conditions in Experiment 1. In both the nonrepeating / no-individuation condition and in the repeating / no-individuation condition, subjects exhibited frequency dominance. In each condition, subjects choose 0000 transfer exemplars more often than 1111 transfer exemplars (nonrepeating $F(1, 30) = 94.59$, $MS_e = .24$, $p < .01$; repeating $F(1, 30) = 28.90$, $MS_e = .24$, $p < .01$). On the typicality test, subjects in both groups again showed a strong preference for the 0000 transfer exemplars (nonrepeating $F(1, 30) = 130.54$, $MS_e = 1.36$, $p < .01$; repeating $F(1, 30) = 49.03$,

TABLE 3
Choice Proportions, Typicality Rankings, Model Fits, and Frequency Estimates for Experiment 2

Measure/Factor	Nonrepeating	Repeating/Individuation None	Number	Side markings	Names	All
Choice proportions						
1111 Transfer exemplars	.05	.26	.24	.11	.31	.03
0000 Transfer exemplars	.94	.74	.84	.88	.69	.96
Typicality rankings						
1111 Transfer exemplars	7.71	6.85	6.63	7.54	6.41	7.84
0000 Transfer exemplars	3.29	4.14	4.38	3.46	4.59	3.14
Choice model fits						
Unweighted linear model	−.96	−.98	−.98	−.97	−.87	−.95
Weighted linear model	.96	.98	.98	.97	.87	.95
Typicality model fits						
Unweighted linear model	−.98	−.99	−.94	−.97	−.67	−1.00
Weighted linear model	.98	.99	.94	.97	.67	1.00
Frequency estimates						
Events	35.31	36.56	33.75	33.38	30.19	29.06
Individuals	—	7.69	5.00	6.75	6.38	5.06

$MS_e = 1.36$, $p < .01$). The results for these two conditions replicate the unequal-frequency conditions in Experiment 1. Subjects showed strong frequency dominance, regardless of whether they received repeating or nonrepeating instructions.

The addition of individuating cues failed to diminish frequency dominance. Across all five repeating conditions, subjects exhibited a strong preference for 0000 transfer exemplars on both the choice task ($F(1, 75) = 249.56$, $MS_e = .24$, $p < .01$) and the typicality task ($F(1, 75) = 243.64$, $MS_e = 1.59$, $p < .01$). Level of individuating information interacted with the degree of preference for the 0000 transfer exemplars, both for choice ($F(4, 75) = 5.44$, $MS_e = .24$, $p < .01$) and typicality ($F(4, 75) = 7.56$, $MS_e = 1.59$, $p < .01$). However, the direction of this interaction was the opposite of that predicted. No condition that provided individuating information significantly decreased the preference for 0000 transfer exemplars relative to the repeating / no-individuation baseline. Instead, the addition of individuating cues increased preference for 0000 transfer exemplars in some conditions, and significantly in the all-individuation condition (choice $F(1, 75) = 5.21$, $MS_e = .24$, $p < .05$; typicality $F(1, 75) = 5.03$, $MS_e = 1.59$, $p < .05$).

Clearly, the presence of individuating information failed to diminish frequency dominance and even increased it in some conditions. Regardless of the type or amount of individuating information received, subjects preferred transfer exemplars similar to one frequent individual more than transfer exemplars similar to four less-frequent individuals. Even when subjects had extensive cues for tracking individuals, they based their categorizations on events rather than on individuals.

Model-based analysis of the data. Fits of the weighted and unweighted linear models confirm the conclusions reached thus far. As Table 3 illustrates, the weighted linear model fit the choice and typicality data much better than the unweighted linear model in all six conditions. These fits confirm the conclusion that subjects weighted Exemplar 5 more heavily than Exemplars 1, 2, 3, and 4.

Analysis of individual subjects. Again, the performance of individual subjects in the repeating conditions suggests caution in drawing a hasty conclusion. In the repeating / no-individuation condition, 11 subjects used events (i.e., preferred 0000 transfer exemplars), 2 subjects used individuals (i.e., preferred 1111 transfer exemplars), and 3 subjects showed no preference. In the number condition, 14 subjects used events, and 2 showed no preference. In the side markings condition, all 16 subjects used events. In the names condition, 11 subjects used events, 3 used individuals, and 2 showed no preference. In the all-individuation condition, all 16 subjects used events. Consistent with the group data, the large majority of repeating subjects used events. Critically, however, five repeating subjects used individuals. Similar to Experiment 1, this finding indicates that the human cognitive system is clearly capable of basing categorization on individuals as well as events.

Frequency estimates. As in Experiment 1, subjects' frequency estimates were reasonably accurate. For events, the average estimates approximated the correct value of 30; for individuals, the average estimates approximated the correct value of 5. Subjects in the name and all-individuation conditions knew the correct number of individuals from the instructions, but subjects in the other three repeating conditions did not. These latter subjects slightly overestimated the number of individuals, but their overestimates were not consistently related to the amount of frequency dominance. Thus, the slight overestimates of individuals in Experiments 1 and 2 do not appear to pose interpretive problems.

The frequency estimates for individuals provide another indication that repeating subjects represented individuals. Repeating subjects who hadn't been told the number of individuals produced estimates near the correct number and far from the total number of events. Indeed, their estimates only differed marginally from those for subjects who had been told the correct number ($F(4, 75) = 2.42$, $MS_e = 8.74$, $p < .10$), suggesting that repeating subjects tracked individuals to a considerable extent.

INDIVIDUALS SAMPLING MODELS

Experiments 1 and 2 demonstrate that human categorization relies heavily on events. Overall, subjects preferred transfer exemplars similar to one frequent individual more than transfer exemplars similar to four less frequent individuals. Even when subjects had many tools for tracking individuals, they continued to exhibit frequency dominance. At first blush, these results confirm events models and disconfirm individuals models. However, additional results implicate individuals in subjects' performance. Repeating subjects were generally quite accurate in estimating the number of individuals, suggesting that they tracked individuals during learning. Furthermore, a few subjects in each experiment based their transfer performance on individuals. Events models cannot account for either result, because they lack mechanisms for tracking individuals and representing them.

This pattern suggests that hybrid models are necessary to explain human categorization. Models that only use events or individuals are insufficient. The space of hybrid models is large. However, two intuitions led us to focus on the class of *individuals sampling models.* The first intuition reflects the fundamental importance of *availability* in human cognition (Tversky & Kahneman, 1973). Across diverse cognitive tasks, some information relevant to a task is more available during retrieval than other relevant information, with the available information controlling performance. Not only does availability enter into decision making (e.g., Kahneman, Slovic, & Tversky, 1982), it also enters into learning (e.g., Estes, 1959), perception (e.g., Bruner & Potter, 1964), categorization (e.g., Nosofsky & Palmeri, 1997), memory (e.g., Rundus, 1973), syntactic parsing (e.g., Fodor, Bever, & Garrett, 1974), reasoning (e.g., Ceraso & Provitera, 1971), problem solving (e.g., Luchins, 1942), and social cognition (e.g., Ross & Nisbett, 1980). In virtually every cognitive activity, some relevant information is more available than other relevant information, with the available information controlling performance. For this reason, incorporating availability into a hybrid model of categorization seemed prudent. As we will see, it allows an individuals model to optimize events naturally and elegantly.

The second intuition that led us to individuals sampling models is the ubiquity of generate-test mechanisms in human cognition. Generate-test mechanisms are closely related to availability. By definition, availability implies that the search for task-relevant information is not exhaustive—only a subset of the relevant information is retrieved. Because the sampling of task-relevant information is partial, it is important to evaluate retrieved information to ensure that it is truly relevant for the task. On some occasions, the most available information may not be appropriate. By coupling a monitoring mechanism to the sampling mechanism, quality control on the sampling process is enforced. If the most available information is not appropriate, the sampling process can iterate until the monitoring process identifies appropriate information. For this reason, cognitive scientists have postulated generate-test mechanisms across the spectrum of cognitive activities. Generate-test mechanisms have been proposed in object recognition (e.g., Bruner & Potter, 1964), memory (e.g., Anderson & Bower, 1972), lexical access (e.g.,

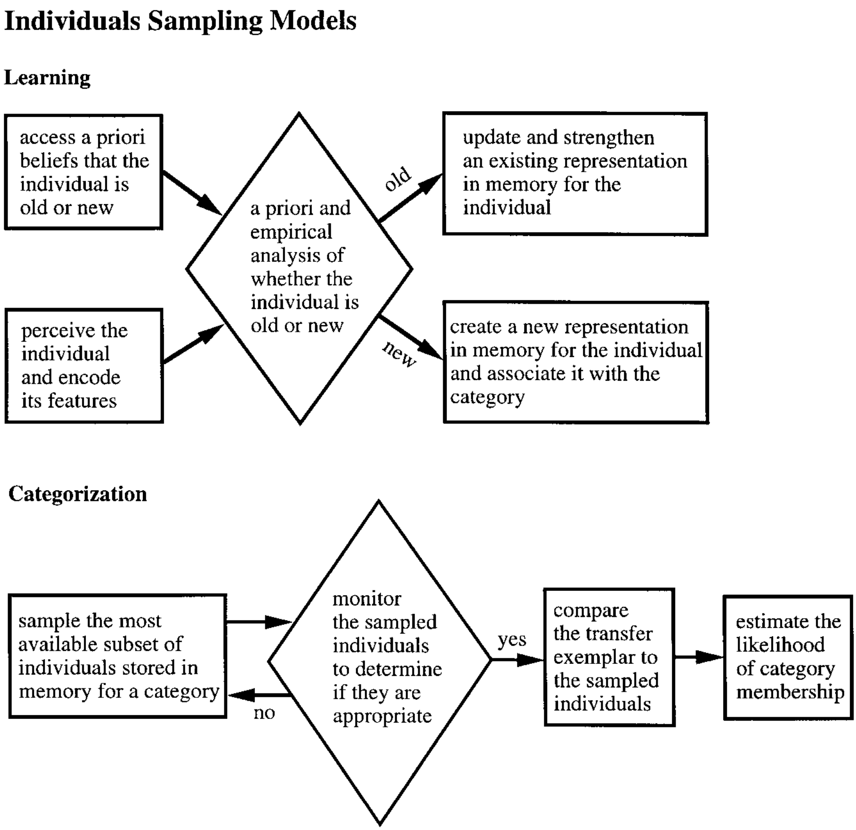


FIG. 5. Learning and categorization assumptions of individuals sampling models.

Swinney, 1979), syntactic parsing (e.g., Fodor et al., 1974), reasoning (e.g., Johnson-Laird, 1983), and problem solving (e.g., Newell & Simon, 1972). In virtually every cognitive activity, evaluating sampled information is essential for satisfactory performance. For this reason, coupling evaluation with availability seemed advantageous.

Figure 5 illustrates the implementation of sampling and monitoring in individuals sampling models. A comparison of Fig. 5 to Fig. 2b illustrates that these models are identical to individuals models during learning. In both classes, a priori and empirical factors determine whether an individual is familiar or unfamiliar, with frames being updated for familiar individuals, and new frames being created for unfamiliar individiuals. Where these two classes of models differ is in categorization. Whereas individuals models (and events models) retrieve category knowledge *exhaustively* for every categorization, individuals sampling models retrieve a *partial subset* of category knowledge and monitor it for appropriateness (Fig. 5). Because retrieving all known individuals for a category exceeds cognitive resources, and because a representative subset typically suffices for satisfactory performance, only the most available subset is sampled initially. Again, the literatures on availability and generate-test mechanisms support this assumption.

To offset the potential risks of partial sampling, the individuals retrieved initially are monitored to ensure their appropriateness for the current categorization. If they are not appropriate, the sampling process iterates until appropriate individuals are found, or until sampling terminates for some other reason. Depending on the situation, different criteria for appropriateness may apply. In a moment, we will consider one criterion that may often be relevant, namely, the representativeness of the sampled individuals. If these individuals are not representative of the category, sampling continues. Once representative individuals are retrieved, they determine the categorization decision.

Importantly, all retrieved individuals are weighted equally at this point, regardless of how often they have occurred previously. If one frequent individual and one infrequent individual are retrieved, they have equal impact on categorization. As discussed earlier for individuals models, if one wants to optimize inferences about the features of individuals, one must weight individuals equally. In the General Discussion, we consider models that weight individuals unequally at this stage of processing.

Individuals sampling models readily handle the primary results of Experiments 1 and 2. Most importantly, they explain most subjects' preference for 0000 transfer exemplars on both the choice and typicality tasks. Because Exemplar 5 is processed more frequently than Exemplars 1, 2, 3, and 4, its frame becomes better established in memory and therefore more available. Frequent processing of Exemplar 5 may add new features, strengthen repeated features, strengthen relations between features, and connect features to other information in memory. Together, these structural factors cause Exemplar 5 to become highly available such that it dominates sampling. Indeed, it becomes so much more available than the other four individuals that it is typically the first and only individual retrieved initially.

A critical issue is why the monitoring process fails to initiate further sampling for more individuals. As discussed shortly, Exemplar 5 may be sufficiently representative of the category that subjects accept it as a basis for categorization. Thus, subjects base categorization on individuals, but only use one. However, subjects also base categorization implicitly on events, because the high frequency of Exemplar 5 causes it to dominate sampling and the subsequent categorization decision.

Individuals sampling models naturally account for individual differences in categorization. According to these models, events typically control categorization because most subjects don't monitor the sampling process carefully. They only sample the most available individual (Exemplar 5), perform minimal evaluation, and accept it as representative of the category. In contrast, a few subjects monitor the sampling process more carefully. On intially retrieving Exemplar 5, they realize that it's only one of several relevant individuals and sample iteratively. After retrieving additional individuals, these subjects proceed to the decision phase, where the larger number of infrequent individuals produces a preference for 1111 transfer exemplars. Thus, individuals sampling models explain individual differences in categorization as individual differences in monitoring.

Finally, individuals sampling models readily explain subjects' accurate estimates of individual frequency. Because subjects update frames for familiar individuals and only create new frames for unfamiliar individuals, they follow the one-entity one-frame principle. As a result, subjects estimate the number of individuals by estimating the number of frames. To estimate event frequency, subjects further consider the amount of information stored in each frame. The more often a frame is updated across repetitions of an individual, the more information it contains. By taking into account both the number of frames in memory for a category, plus the amount of information stored in each, subjects estimate event frequency. In the General Discussion, we present a specific proposal of this process (Fig. 6).

Thus far, the empirical evidence for individuals sampling models comes from post hoc reinterpretations of Experiments 1 and 2. Although individuals sampling models readily explain the entire pattern of results in these experiments, it is essential to evaluate these models a priori. The next two experiments serve this purpose. Experiments 3 and 4 both test the prediction that subjects monitor a partial sampling process during categorization. Both experiments also test the predictions of all individuals models that subjects establish one frame for each individual during learning, and then use these frames later to make categorization decisions.

EXPERIMENT 3

How might one test the prediction that coupled sampling and monitoring processes underlie categorization? The literature on stereotypes provides a suggestion. To maintain a stereotype, subjects often discount unusual individuals (e.g., Kunda & Oleson, 1995, 1997). For example, encountering an ethical politician may not change one's stereotype of politicians as unethical, because the individual is perceived as so unusual as to be uninformative about the category. When an unusual individual violates a stereotype, its potential impact on learning is often suppressed.

This finding suggests that Exemplar 5 in Experiments 1 and 2 was not sufficiently unusual to be discounted during monitoring as a category member. If Exemplar 5 had been sufficiently unusual, subjects should have discounted it, at least that's what the stereotypes literature suggests. Instead, subjects may have perceived Exemplar 5 as lying within the acceptable range of variability for a category of fish, given individuals in natural categories vary idiosyncratically (Fried & Holyoak, 1984; Nisbett & Kunda, 1985).

When only Exemplar 5 came to mind initially because of its high availability, subjects may not have thought twice about using it, such that it dominated performance.[9] Exemplar 5's high availability may also produce an experience of fluency that endows it with further credibility as a representative category member (Jacoby, Kelley, Brown, & Jasechko, 1989).

This line of reasoning suggests that if Exemplar 5 were distanced further from Exemplars 1, 2, 3, and 4, it might become unusual enough to lie outside the category's range of acceptable variability. If so, then subjects might discount it during monitoring, even if it's high availability causes it to be sampled initially. As a result, subjects sample further and retrieve less frequent individuals that eventually control categorization. To test this prediction, Experiment 3 included three groups of subjects: the repeating / close-similarity condition, the repeating / far-similarity condition, and the nonrepeating / far-similarity condition. The repeating / close-similarity condition was a replication of the repeating / all-individuation condition of Experiment 2. ''Close-similarity'' means that Exemplar 5 was relatively close to Exemplars 1, 2, 3, and 4, as defined in Table 1.

The other two conditions received the far-similarity materials in Table 4, one under repeating instructions (with all of the individuating cues from Experiment 2), and the other under nonrepeating instructions. As a comparison of Tables 1 and 4 illustrates, Exemplars 1, 2, 3, and 4 are common to both the close- and far-similarity materials. These materials only differ in two ways. First, Exemplar 5′ is much less similar to Exemplars 1, 2, 3, and 4 in the far-similarity materials than is Exemplar 5 to Exemplars 1, 2, 3, and 4 in the close-similarity materials. The concrete differences between Exemplars 5 and 5′ can be seen in Fig. 3. Second, Exemplars 11′, 12′, 13′, 14′, and 15′ in the far-similarity materials are much less similar to their pair members on the transfer test than Exemplars 11, 12, 13, 14, and 15 are to theirs, although the similarity of Exemplars 11′, 12′, 13′, 14′, and 15′ to Exemplar 5′ is the same as the similarity of Exemplars 11, 12, 13, 14, and 15 to Exemplar 5. In all three conditions, Exemplars 1, 2, 3, and 4 were presented 3 times each, and Exemplar 5 or 5′ was presented 18 times.

Two planned comparisons between these three groups allow us to test predictions of independent sampling models. The first contrasts the two repeating conditions to test the prediction that coupled sampling and monitoring processes underlie categorization. Individuals sampling models predict that repeating subjects should prefer 0000 transfer exemplars in the close-similarity condition but prefer 1111 transfer exemplars in the far-similarity condition. In the close-similarity condition, Exemplar 5 should dominate sampling initially because of its high availability. It should then pass monitoring because it lies within the category's range of acceptable variability (as suggested by Experiments 1 and 2). In the far-similarity condition, Exem-

[9] Patrick Shrout suggested this explanation.

TABLE 4
Category Structure in the Far-Similarity Condition of Experiment 3

Exemplar	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Presentation frequency
	Acquisition exemplars										
1	1	1	1	1	1	1	1	1	1	1	3
2	2	2	1	1	1	1	1	1	1	1	3
3	3	3	1	1	1	1	1	1	1	1	3
4	4	4	1	1	1	1	1	1	1	1	3
5′	5	5	0	0	0	0	0	0	0	0	18

Exemplar	D1	D2[a]	D3	D4	D5	D6	D7	D8	D9	D10	Contrast exemplar
	Transfer exemplars										
6	6	6/7	1	1	1	1	1	1	1	1	11
7	6	6/7	1	1	1	0	1	1	1	1	12
8	6	6/7	1	1	0	1	1	1	1	1	13
9	6	6/7	1	0	1	1	1	1	1	1	14
10	6	6/7	0	1	1	1	1	1	1	1	15
11′	6	6/7	0	0	0	0	0	0	0	0	6
12′	6	6/7	0	0	0	1	0	0	0	0	7
13′	6	6/7	0	0	1	0	0	0	0	0	8
14′	6	6/7	0	1	0	0	0	0	0	0	9
15′	6	6/7	1	0	0	0	0	0	0	0	10

[a] The two exemplars in each transfer pair always had the same value on D2, either 6 or 7.

plar 5′ should again dominate sampling because of its high availability. However, it should then produce a negative monitoring response, because it was characterized during learning as lying outside the acceptable range of variability. As a result, the likelihood of sampling of Exemplars 1, 2, 3, and 4 increases, and subjects are more likely to exhibit a preference for 1111 transfer exemplars.

A second planned comparison tests the predictions of all individuals models that subjects establish one frame for each individual during learning, and then use these frames later to make categorization decisions. Of interest here are the two far-similarity conditions. Subjects in both conditions received the same sequence of training exemplars, but one group studied them under repeating instructions, whereas the other group studied them under nonrepeating instructions. Individuals sampling models predict that these two groups should contrast in transfer performance. Repeating subjects should prefer 1111 transfer exemplars, because they discount the early sampling of Exemplar 5′ and sample Exemplars 1, 2, 3, and 4. In contrast, nonrepeating subjects should prefer 0000 transfer exemplars. Under nonrepeating instructions, subjects should believe that there are 30 individuals, 18 taking the form of Exemplar 5′, and 12 total taking the form of Exemplars 1, 2, 3, and 4. Because individuals taking the form of Exemplar 5′ are most numerous, they dominate sampling and pass monitoring, seeming highly representative during learning because of their numerosity. This prediction hinges on the frame creation process. In the repeating condition, subjects should only construct 5 frames, such that the frame for Exemplar 5′ is not representative. In the nonrepeating condition, however, subjects should construct 30 frames, such that the 18 frames for Exemplar 5′ are representative.[10]

Method

Design and subjects. One between-subject variable and one within-subject variable were crossed to structure the experiment. The between-subject variable contrasted three groups of subjects: repeating / close similarity, repeating / far similarity, and nonrepeating / far similarity. The within-subject variable contrasted 1111 transfer exemplars with 0000 transfer exemplars, as in the previous experiments. The dependent measures were again categorization choices, frequency estimates, and typicality rankings. Subjects were 48 members of the University of Chicago community, who received $2.00 for 15 to 20 minutes of participation. Sixteen subjects were assigned randomly to each of the three between-subject conditions. The different versions of the training and test materials were distributed equally across the experimental variables.

Materials. Subjects in the repeating / close-similarity condition received the same materials as the repeating / all-individuation subjects in Experiment 2. The training and transfer materials were those in Table 1, and subjects received names for exemplars during training only.

Subjects in the two far-similarity conditions received the materials in Table 4. In the repeating / far-similarity condition, subjects received names for exemplars during training. In the nonrepeating / far-similarity condition, subjects received no names, given the nonrepeating instructions made individuating cues unnecessary. All other aspects of the training materials were the same as in the previous experiments. The transfer materials were identical as well, except that Exemplars 11′, 12′, 13′, 14′, and 15′ replaced Exemplars 11, 12, 13, 14, and 15.

Procedure. To maximize subjects' ability to track individuals during training, subjects in the two repeating conditions received the same training instructions as the all-individuation subjects in Experiment 2 (i.e., they were told about number, side markings, and names). Subjects in the nonrepeating condition received the same nonrepeating instructions as nonrepeating subjects in Experiments 1 and 2. All other aspects of the procedure were the same as in the previous experiments.

Results

Choice proportions and typicality rankings. Two planned comparisons assess the predictions of individuals sampling models. According to the first, subjects in the repeating / far-similarity condition should exhibit a preference for 1111 transfer exemplars, because Exemplar 5′ produces a negative moni-

[10] Because both planned comparisons include the repeating / far condition, one might worry that they are not independent. However, we could have performed a single planned comparison across all three conditions simultaneously that would have used the repeating / far condition only once, and the results would have supported the same conclusions. We organize analyses of the three conditions around two planned comparisons to assess the a priori hypotheses of interest, and to facilitate their presentation.

TABLE 5
Choice Proportions, Typicality Rankings, Model Fits, and Frequency Estimates for Experiment 3

Measure/Factor	Close similarity: Repeating	Far similarity: Repeating	Far similarity: Nonrepeating
Choice proportions			
1111 Transfer exemplars	.03	.69	.00
0000 Transfer exemplars	.95	.31	1.00
Typicality rankings			
1111 Transfer exemplars	7.69	4.59	7.99
0000 Transfer exemplars	3.31	6.41	3.01
Choice model fits			
Unweighted linear model	−.96	1.00	−1.00
Weighted linear model	.96	−1.00	1.00
Typicality model fits			
Unweighted linear model	−.94	.90	−.97
Weighted linear model	.94	−.90	.97
Frequency estimates			
Events	33.00	33.38	29.06
Individuals	5.00	4.94	—

toring response that initiates further sampling. In contrast, subjects in the repeating / close-similarity should exhibit a preference for 0000 transfer exemplars, because Exemplar 5 does not produce a negative monitoring response. As Table 5 illustrates, the results support this prediction. The close-similarity condition replicates the all-individuation condition in Experiment 2, with subjects preferring 0000 transfer exemplars on the choice task ($F(1, 30) = 32.44$ $MS_e = .75$, $p < .01$). The repeating / far-similarity condition shows the opposite effect, with these subjects preferring 1111 transfer exemplars ($F(1, 30) = 5.20$, $MS_e = .75$, $p < .05$). All 16 subjects in the close-similarity condition preferred 0000 transfer exemplars, whereas 11 of the 16 subjects in the repeating / far-similarity condition preferred 1111 transfer exemplars, and 5 preferred 0000 transfer exemplars. The interaction between similarity and transfer exemplar was significant ($F(1, 30) = 31.82$, $MS_e = .75$, $p < .01$), indicating that the similarity manipulation modulated the performance of repeating subjects.

The typicality rankings exhibited the same pattern. In the repeating / close-similarity condition, subjects preferred 0000 transfer exemplars ($F(1, 30) = 27.67$, $MS_e = 5.53$, $p < .01$); in the repeating / far similarity condition, subjects preferred 1111 transfer exemplars ($F(1, 30) = 4.82$, $MS_e = 5.53$, $p < .05$). Again, similarity and transfer exemplar interacted ($F(1, 30) = 27.79$, $MS_e = 5.53$, $p < .01$).

When a highly unusual individual was present in the repeating / far-similarity condition, most subjects did not allow it to dominate categorization, even though it was probably highly available. This finding suggests that subjects established representations of individuals during learning and monitored them during sampling. This finding further suggests that subjects in the close-similarity condition did not perceive Exemplar 5 as unusual, given all of them allowed it to dominate categorization. Interestingly, 5 of the 16 subjects in the repeating / far-similarity condition similarly allowed Exemplar 5′ to dominate, underlining the potency of availability in sampling.

The second planned comparison concerns the two far-similarity conditions. Both groups received the same sequence of training exemplars. However, the repeating / far-similarity group believed that 5 individuals appeared on multiple occasions, whereas the nonrepeating / far-similarity group believed that 30 different individuals each appeared once. As discussed earlier, individuals sampling models predict that subjects in the repeating / far-similarity condition should prefer 1111 transfer exemplars, because the four individuals for Exemplars 1, 2, 3, and 4 dominate the one individual for Exemplar 5′. In contrast, subjects in the non-repeating / far-similarity condition should prefer 0000 transfer exemplars, because the 18 individuals for Exemplars 5′ dominate the 12 individuals for Exemplars 1, 2, 3, and 4.

The results in Table 5 support this prediction of individuals sampling models. Varying the instructions that subjects received about individuals reversed their categorization preferences. Although both groups received the same sequence of training exemplars, their different beliefs about individuals produced substantially different performance. On the choice task, repeating / far-similarity subjects preferred 1111 transfer exemplars ($F(1, 30) = 5.32$ $MS_e = .73$, $p < .05$), whereas nonrepeating / far-similarity subjects preferred 0000 transfer exemplars ($F(1, 30) = 37.72$, $MS_e = .73$, $p < .01$). All 16 subjects in the nonrepeating condition exhibited a 0000 preference. The interaction between instructions and transfer exemplar was significant ($F(1, 30) = 35.69$, $MS_e = .73$, $p < .01$), indicating that the differing instructions about individuals modulated performance.

The typicality rankings exhibited the same pattern. In the repeating / far-similarity condition, subjects preferred 1111 transfer exemplars ($F(1, 30) = 4.90$, $MS_e = 5.43$, $p < .05$); in the non-repeating / far similarity condition, subjects preferred 0000 transfer exemplars ($F(1, 30) = 36.46$, $MS_e = 5.43$, $p < .01$). Again, instructions and transfer exemplar interacted ($F(1, 30) = 34.05$, $MS_e = 5.43$, $p < .01$).

These results indicate that repeating subjects constructed frames for 5 individuals, whereas non-repeating subjects constructed frames for 30 individuals. In the context of 5 individuals, Exemplar 5′ appeared unusual, such that repeating subjects sampled other exemplars that ultimately dominated categorization. In the context of 30 individuals, the 18 individuals for Exemplar 5′ appeared representative, such that nonrepeating subjects allowed them to dominate instead.

Model-based analysis of the data. Fits of the weighted and unweighted linear models confirm the conclusions reached thus far. As Table 5 illustrates, the weighted linear model best fit the choice and typicality data in the repeating / close-similarity condition and the nonrepeating / far-similarity condition. In the repeating / far-similarity condition, however, the unweighted linear model provided the best fit. This pattern confirms that Exemplar 5 was weighted more than Exemplars 1, 2, 3, and 4 in the repeating / close-similarity condition and in the nonrepeating / far-similarity condition, whereas the five exemplars were weighted more equally in the repeating / far-similarity condition.[11]

Frequency estimates. As in Experiments 1 and 2, subjects' frequency estimates were reasonably accurate (see Table 5). For events, the average estimates approximated the correct value of 30. For individuals, subjects' estimates are not of much interest, given all repeating subjects were told the number of individuals during acquisition.

Discussion

These results support the a priori prediction of individuals sampling models that categorization includes coupled sampling and monitoring stages. In the repeating / close-similarity condition, Exemplar 5 was again so highly available that it dominated sampling. Because it was sufficiently similar to Exemplars 1, 2, 3, and 4, it passed monitoring and dominated the categorization decision. In contrast, when Exemplar 5′ was sufficiently distanced from Exemplars 1, 2, 3, and 4 in the repeating / far-similarity condition, it failed the monitoring stage, such that the further sampling of other individuals ultimately dominated categorization. Nevertheless, the availability of Exemplar 5′ was still sufficiently high that 5 of the 16 subjects allowed it to dominate their decisions. Presumably, these subjects monitored the sampling process less carefully than the other 11 subjects.

These findings pose problems for events models that store independent memories of a repeated individual. Without mechanisms for integrating the memories of an individual across events, these models cannot explain performance in the repeating / far-similarity condition. If the 18 presentations for Exemplar 5′ aren't integrated, it can't appear unrepresentative in a contrast with the other 4 individuals, such that they are sampled and control categorization. Instead, Exemplar 5′ should always dominate, as in the nonrepeating / far-similarity condition, where subjects presumably represented 18 individuals that took the form of Exemplar 5′. These results indicate that

[11] The interpretation of the weighted linear model in Experiments 3 and 4 is more general than its interpretation in Experiments 1 and 2. In Experiments 1 and 2, a good fit of the weighted model was interpreted as indicating that frequent exemplars established more traces in memory than infrequent exemplars. In Experiments 3 and 4, a good fit of this model is also interpreted as indicating that a frequent individual established a more available frame in memory than infrequent exemplars. Thus, a good fit of the weighted linear model is consistent with either events model or individual sampling models. If frequency didn't bias sampling, only the unweighted linear model should provide a good fit.

a mechanism for integrating information about individuals across events is essential to a satisfactory theory of categorization.

Experiment 3 rules out two additional accounts of frequency dominance that we haven't considered thus far. According to one, the most frequent individual dominates categorization because its features are most reliable.[12] Because subjects are more certain about the features of a frequent individual than those of an infrequent individual, they allow the frequent individual to have more impact on categorization. However, the data in Experiment 3 contradict this hypothesis. If reliability were the reason for frequency dominance, then Exemplar 5′ should have dominated categorization in the repeating / far-similarity condition. Because Exemplar 5′ occurred more often than the other exemplars, its features were more reliable. As we saw, however, these subjects allowed the four less reliable exemplars to dominate. Rather than reliability being the important factor, availability was critical.

The results in the nonrepeating/ far-similarity condition also rule out the account that subjects in the repeating / far-similarity condition found Exemplar 5′ so unappealing or weird that they developed a negative response bias toward it. On this view, the repeating / far-similarity condition preferred 1111 transfer exemplars, not because they were close to four similar individuals, but because they were dissimilar to a disliked training exemplar. If this account were correct, the nonrepeating/ far-similarity condition should have also preferred 1111 transfer exemplars. Because all 16 of these subjects strongly preferred 0000 transfer exemplars, subjects did not appear to stigmatize Exemplar 5′.

Finally, it is important to consider the process by which Exemplar 5′ comes to be viewed as unrepresentative. During the training phase, subjects may have noticed that Exemplar 5′ was different from the other four individuals and updated its frame with this observation. When subjects sampled Exemplar 5′ at transfer, they may have retrieved this information, thereby producing a negative monitoring response and further sampling. We explore this process in the General Discussion.

EXPERIMENT 4

The final experiment provides two further tests of the hypothesis that coupled sampling and monitoring processes underlie categorization. According to individuals sampling models, frequency dominance occurs because subjects stop sampling after retrieving Exemplar 5, which they perceive as representative. If subjects' failure to sample Exemplars 1, 2, 3, and 4 produces their preference for 0000 transfer exemplars, then making subjects aware of these training exemplars during transfer should increase their impact. To increase the sampling of these exemplars, subjects in the repeating / cues

[12] Gerd Gigerenzer suggested this hypothesis.

condition received the names of all five training individuals as they performed the transfer tasks. In the choice task, these names appeared at the top of each page that presented a pair of transfer exemplars. In the typicality task, these names appeared on the page that preceded the 10 transfer exemplars for ranking. If frequency dominance occurred in the previous experiments because subjects only considered Exemplar 5, it should decrease when cues increase the availability of Exemplars 1, 2, 3, and 4.

Whereas the repeating / cues condition attempted to reduce frequency dominance by increasing the breadth of sampling, the repeating / protocols condition attempted to reduce frequency dominance by increasing the amount of monitoring. These latter subjects produced protocols as they performed the standard version of the experiment with no cues at transfer. Because protocol production is inherently a monitoring process, it should increase how much subjects monitor the sampling stage of categorization. As a result, subjects should be more likely to perceive Exemplar 5 as unusual and to sample further individuals, thereby increasing the likelihood of a 1111 preference. Other researchers have similarly found that increasing the level of monitoring causes subjects to retrieve additional information that improves performance (e.g., Greenwald & Banaji, 1995; Smith & Sloman, 1994).

Experiment 4 also tested the predictions of all individuals models that subjects track individuals during learning, and that they consider individuals during categorization. To test these predictions, the extent to which subjects discussed individuals in the repeating / protocols condition was assessed.

Finally, a third condition without cues or protocols established a baseline against which to assess the predictions of the repeating / cues condition and the repeating / protocols condition. In all three conditions, subjects received the close-similarity materials of the previous experiments under repeating instructions, with Exemplar 5 presented more often than Exemplars 1, 2, 3, and 4. All three sources of individuating information from Experiment 2 were included to ensure optimal individuation.[13]

Method

Design and subjects. One between-subject variable and one within-subject variable were crossed to structure the experiment. The between-subject variable contrasted three groups of subjects: repeating / cues, repeating / protocols, and repeating / no cues or protocols. The within-subject variable contrasted 1111 transfer exemplars with 0000 transfer exemplars, as in the previous experiments. The dependent measures were again categorization choices, frequency estimates, and typicality rankings. Subjects were 44 members of the University of

[13] Because both planned comparisons use the same baseline condition, one might worry that they are not independent. As in Experiment 3, however, we could have performed a single planned comparison across all three conditions simultaneously that would have used the baseline condition only once, and the results would have supported the same conclusions. Again, we organize analyses of the three conditions around these two comparisons to reflect the a priori hypotheses of interest, and to facilitate their presentation.

Chicago community, who received $2.00 for 15 to 20 minutes of participation. Subjects were assigned randomly to the three between-subject conditions, with 16 in the cues condition, 12 in the protocols condition, and 16 in the baseline condition. The different versions of the training and test materials were distributed equally across the experimental variables.

Materials. The training materials for all three conditions were identical to those in the repeating / all-individuation condition of Experiment 2, and in the repeating / close-similarity condition of Experiment 3. All groups received the close-similarity materials from previous experiments with names for the individuals. Again, Exemplars 1, 2, 3, and 4 were presented 3 times each, and Exemplar 5 was presented 18 times.

The transfer materials for all three conditions were the same as in these earlier conditions with the following exceptions. In the repeating / cues condition, each 21.5 by 28 cm page of the choice materials began with the statement, ''Here are the names of the five individuals from the fish species that you saw earlier,'' followed below by the names in a row. Under the names was the statement, ''Put an X under the fish below that is more likely to belong to the same species,'' followed by the two transfer exemplars, each 5 cm tall. For the typicality materials, a cover page stated, ''Here are the names of the five individuals from the fish species that you saw earlier,'' followed below by the names in a row. Under the five training individuals was the statement, ''On the next page, rank the fish for how good an example each is of the same species,'' followed by the instructions described earlier for rankings. On the next page, the 10 transfer exemplars appeared in a random order, each 6 cm tall. On each page of a subject's test booklet, the names of the five training individuals appeared in the same random order, with half of the subjects receiving them in one random order, and the other half receiving them in another. Subjects in the protocols and no-cues conditions received the same test booklets as subjects in the cues condition but without the names of the training exemplars and the statements preceding them.

Procedure. All aspects of the training and transfer procedures were identical to those in the all-individuation condition of Experiment 2 and the repeating / close-similarity condition of Experiment 3 with two exceptions. First, some of the transfer instructions were presented in the test booklets rather than verbally by the experimenter. Second, in the protocols condition, subjects were tape recorded as they described ''what they were thinking'' during each learning and test trial.

Results

Choice proportions and typicality rankings. Two planned comparisons assess the predictions of individuals sampling models for the cues and protocols manipulations. In the first, individuals sampling models predict that cues should weaken subjects' preference for 0000 transfer exemplars by making less established individuals more available. As Table 6 illustrates, the results support this prediction. When subjects received no cues at transfer, strong frequency dominance again occurred. As Table 6 illustrates, these subjects overwhelmingly preferred the 0000 transfer exemplars over the 1111 transfer exemplars on both the choice task ($F(1, 30) = 110.73$, $MS_e = .23$, $p < .01$) and the typicality task, ($F(1, 30) = 79.05$, $MS_e = 1.79$, $p < .01$). When subjects received cues, however, frequency dominance diminished, as individuals sampling models predict. A preference for 0000 transfer exemplars remained evident on both the choice task ($F(1, 30) = 46.79$, $MS_e = .65$, $p < .01$) and the typicality task ($F(1, 30) = 25.93$, $MS_e = 4.84$, $p < .01$). More importantly, however, the presence of cues diminished the size of this preference relative to subjects who didn't receive cues, as indicated by sig-

TABLE 6
Choice Proportions, Typicality Rankings, Model Fits, and Frequency Estimates for Experiment 4

Measure/Factor	No cues or protocols	Cues	Protocols
Choice proportions			
1111 Transfer exemplars	.01	.24	.63
0000 Transfer exemplars	.98	.76	.42
Typicality Rankings			
1111 Transfer exemplars	7.60	6.20	4.97
0000 Transfer exemplars	3.40	4.80	6.03
Choice model fits			
Unweighted linear model	−.95	−.94	.77
Weighted linear model	.95	.94	−.77
Typicality model fits			
Unweighted linear model	−.97	−.94	.89
Weighted linear model	.97	.94	−.89
Frequency estimates			
Events	29.63	28.44	30.48
Individuals	6.25	5.00	5.08

nificant interactions between cue condition and transfer exemplar (choice $F(1, 30) = 4.22$, $MS_e = .65$, $p < .05$; typicality $F(1, 30) = 6.48$, $MS_e = 6.48$, $p < .05$). Whereas none of the 16 subjects in the no cues condition preferred 1111 over 0000 exemplars, 4 of the 16 subjects in the cues condition did.[14]

Although the name cues reduced frequency dominance, they did not eliminate it. This is not surprising, given subjects were not required to produce each individual's features when given its name during learning. Thus, some subjects may have experienced difficulty during sampling in using the name cues to retrieve the associated individuals. Experiments designed to maximize the retrieval of individuals from names would probably diminish frequency dominance further.

The second planned comparison addresses the effect of protocols on transfer performance. Individuals sampling models predict that protocols should increase monitoring of the sampling process, thereby weakening subjects' preference for 0000 transfer exemplars. As Table 6 illustrates, the results support this prediction. When subjects produced protocols, frequency domi-

[14] The finding that all 16 subjects in the repeating / all-individuation condition of Experiment 4 preferred 0000 transfer exemplars represents a stable phenomenon. As mentioned earlier, all 16 subjects in the repeating / all-individuation condition of Experiment 2 exhibited this preference, as did all 16 subjects in the repeating / all-individuation condition of Experiment 3. Obtaining complete unanimity in all three conditions decreases the probability that four subjects' 1111 preference in the repeating / cues condition represents a random fluctuation in performance. Instead, it appears that cues decrease frequency dominance.

nance diminished, with the preference for 0000 transfer exemplars over 1111 transfer exemplars disappearing on both the choice task ($F(1, 26) = 2.37$, $MS_e = .45$, ns) and the typicality task ($F(1, 26) = 1.75$ $MS_e = 3.89$, ns); in fact, subjects' preference reversed slightly. More importantly, the presence versus absence of protocol production moderated subjects' preference for 0000 versus 1111 transfer exemplars, as indicated by significant interactions between these variables (choice $F(1, 26) = 36.50$, $MS_e = .45$, $p < .01$; typicality $F(1, 26) = 95.10$, $MS_e = 3.89$, $p < .01$). Whereas none of the 16 subjects in the no cue condition preferred 1111 over 0000 transfer exemplars, 6 of the 12 protocol subjects preferred 1111 transfer exemplars, 4 preferred 0000 transfer exemplars, and 2 showed no preference. Thus, producing protocols diminished frequency dominance, as individuals sampling models predict.

Model-based analysis of the data. Fits of the weighted and unweighted linear models confirm the conclusions reached so far. As Table 6 illustrates, the weighted linear model best fit the choice and typicality data in the cues and baseline conditions, with the presence of cues reducing the fit slightly. This pattern confirms that most of these subjects weighted Exemplar 5 more than Exemplars 1, 2, 3, and 4. In contrast, the unweighted linear model best fit the choice and typicality data in the repeating / protocols condition, reflecting these subjects' overall tendency to weight individuals equally.

Frequency estimates. As in the previous experiments, subjects' frequency estimates were reasonably accurate (see Table 6). For events, the average estimates approximated the correct value of 30. For individuals, subjects' estimates are not of much interest, given subjects were told the number of individuals during acquisition.

Learning protocols. For each of the 30 learning trials, a subject's statements were coded for four types of content: (1) Did the subject track individuals across repetitions? (2) What features of individuals did the subject notice? (3) Did the subject compare individuals globally to one another? (4) Did the subject compare individuals on specific features? Appendix C presents examples of statements for each type of content. Analyses were performed separately for trials 1 to 10 and trials 11 to 30. During trials 1 to 10, Exemplars 1, 2, 3, and 4 were each presented for the first time, and Exemplar 5 was presented six times. During trials 11 to 30, Exemplars 1, 2, 3, and 4 were each presented two more times, and Exemplar 5 was presented 12 more times. Thus, trials 1 to 10 represent the period in which subjects first saw all five exemplars, whereas trials 11 to 30 represent the period in which all presentations were repetitions of previously encountered individuals.[15] All ANOVAs for probabilities were performed on arcsin transformed data.

[15] Another way of assessing learning histories is to compare all five exemplars across presentations 1, 2, and 3. The problem with this is that Exemplar 5′s first three presentations occur during the first five trials, such that its presentations 2 and 3 occur very early in learning,

The first analysis addresses whether subjects tracked individuals. According to all individuals models, subjects should track an individual so that they can integrate information across its repetitions into a common frame. To assess this issue, the protocols produced for each training exemplar—from the second presentation on—were coded for the presence of tracking. Two types of statements indicate the presence of this activity. In ''Recognition'' statements, subjects stated unambiguously that they recognized a previously presented individual. As the examples in Appendix C illustrate, these statements clearly indicate that knowledge of an individual stored in memory was activated to process its current presentation. In ''Naming'' statements, subjects simply stated the name of the individual at the beginning of a trial. Although these may have often been statements of recognition, they could also have resulted from subjects simply reading an exemplar's name on a stimulus card. Thus, these statements must be viewed with caution.

As Table 7 illustrates, there is overwhelming evidence that subjects tracked individuals. For Exemplar 5, the probability of explicit recognition was .98 for trials 1 to 10 and .85 for trials 11 to 30. If one assumes that some proportion of explicit namings indicate tracking as well, subjects accessed previously established knowledge even more often. For Exemplars 1, 2, 3, and 4, subjects only had opportunities to exhibit tracking during trials 11 to 30. During this period, subjects produced explicit recognitions much less often than for Exemplar 5, and they produced explicit namings much more often (interaction $F(1, 11) = 12.81$, $MS_e = .37$, $p < .05$). This suggests that subjects were less able to track repetitions of Exemplars 1, 2, 3, and 4 than repetitions of Exemplar 5. This further suggests that subjects' slight overestimates of individual frequency in earlier experiments may have resulted from sporadic failures to track Exemplars 1, 2, 3, and 4, such that another frame for an individual was mistakenly established rather than updating an existing one. Nevertheless, subjects explicitly recognized Exemplars 1, 2, 3, and 4 on at least 57% of the relevant trials, clearly indicating that they were tracking these individuals a majority of the time.

The second analysis examines the features of individuals that subjects noticed. In Table 7, the probability of mentioning a feature on a unique dimension is the probability per unique dimension that, on a given trial, a subject mentioned a feature that individuated the current individual (top fin, side marking). In contrast, the probability of mentioning a feature on a shared dimension is the probability per shared dimension that, on a given trial, a subject mentioned a feature that Exemplars 1, 2, 3, and 4 shared, or a feature that distinguished Exemplar 5 from the other four exemplars (bottom fin, back fin, mouth, face). Appendix C provides examples of these statements.

---

whereas presentations 2 and 3 for the other four exemplars occur much later. Because the protocols seemed to vary more across trials than across presentations, we report the results according to trials. When relevant, comparisons based on presentations are noted.

TABLE 7
Probabilities and Frequencies per Subject per Trial per Opportunity for the Content of the Learning Protocols in Experiment 4

Content category	Exemplars 1, 2, 3, and 4		Exemplar 5	
	Trials 1–10	Trials 11–30	Trials 1–10	Trials 11–30
Probability of tracking an individual				
Recognition	—	.57	.98	.85
Naming	—	.31	.02	.07
Probability of mentioning a feature				
Unique dimension	.77	.61	.31 (.71)[a]	.15 (.29)[b]
Shared dimension	.58	.47	.27 (.42)[a]	.20 (.28)[b]
Frequency of global comparisons				
Self	—	.03	.33	.18
Contrast	.05	.11	.01	.02
Neighbor	.17	.04	—	—
Previous	.05	.05	.11	.06
Frequency of specific comparisons				
Self	—	.05	.32	.17
Contrast	.59	.36	.11	.06
Neighbor	.89	.31	—	—
Previous	.79	.28	.29	.13

Note. For Exemplars 1, 2, 3, and 4, the first presentation occurred during trials 1 to 10, and the second and third presentations occurred during trials 11 to 30. For Exemplar 5, the first six presentations occurred during trials 1 to 10, and the last 12 presentations occurred during trials 11 to 30.

[a] Probability for the first presentation.

[b] Probability for the second and third presentations.

As Table 7 illustrates, the probability of mentioning a feature was higher for Exemplars 1, 2, 3, and 4 than for Exemplar 5, both across all 30 trials (.61. vs .23; $F(1, 11) = 110.64$, $MS_e = .16$, $p < .01$), and when only the first three presentations of all five exemplars are compared (.61. vs. .43; $F(1, 11) = 20.85$, $MS_e = .21$, $p < .01$). A likely explanation is that subjects saw Exemplar 5 so often that they didn't process it as extensively on each occasion as the less frequent exemplars. Subjects may have worked harder at learning Exemplars 1, 2, 3, and 4 because they had fewer opportunities to do so. Across all 30 trials, subjects generally processed features on the unique dimensions more often than features on the shared dimensions ($F(1, 11) = 22.31$, $MS_e = .03$, $p < .01$). However, this preference only held for Exemplars 1, 2, 3, and 4 (.69 versus .43) and not for Exemplar 5 (.22 vs. 23) (interaction $F(1, 11) = 7.71$, $MS_e = .12$, $p < .05$). Subjects may have focused more on the individuating features of Exemplars 1, 2, 3, and 4 because they were so similar to one another. Nevertheless, subjects did process the remaining four dimensions often as they attempted to understand how the five individuals were similar and different.

The third and fourth analyses address how subjects attempted to understand the similarities and differences between individuals. In global comparisons, subjects noted that one individual was similar to or different from another individual (or itself) without stating specific features as the basis of comparison. In specific comparisons, subjects stated specific features. Appendix C provides examples.

For both global and specific comparisons, four types were of interest: self, contrast, neighbor, and previous comparisons. When making a self comparison, subjects went beyond recognizing an individual and compared its current presentation to existing knowledge of the same individual established on earlier presentations. When making contrast comparisons, subjects were studying Exemplar 1, 2, 3, or 4 and drew a comparison to Exemplar 5, or they were studying Exemplar 5 and drew a comparison to Exemplar 1, 2, 3, or 4. These comparisons are contrastive because they usually noted differences between these two subsets of individuals. When making neighbor comparisons, subjects were studying Exemplar 1, 2, 3, or 4, and drew a comparison to another individual in this subset. Finally, when making previous comparisons, subjects compared the current individual under study to the individual studied on the previous trial.

Table 7 presents the average frequency of comparisons per trial from this analysis. These average frequencies are normalized by the number of opportunities for producing them, which often involved detailed analyses of the presentation sequences. For example, the contrast frequencies take into account the fact that, once all individuals had been presented once, four possible contrasts could be drawn from Exemplar 5 to Exemplars 1, 2, 3, and 4, whereas only one could be drawn from Exemplar 1, 2, 3, or 4 to Exemplar 5. Similarly, the neighbor frequencies take into account the fact that Exemplar 1, 2, 3, and 4 each had three possible neighbors. Furthermore, because the number of possible comparisons increased across trials 1 to 10 as each individual occurred for the first time, the average number of possible comparisons was computed trial by trial to normalize the average frequencies.

Subjects produced three times as many specific comparisons per opportunity as global comparisons (1.09 vs .30, $F(1, 11) = 30.04$, $MS_e = .49$, $p < .01$).[16] This suggests that subjects performed more analytic than holistic processing (e.g., Kemler Nelson, 1984; Ward & Scott, 1987). Subjects produced twice as many comparisons per Exemplar 1, 2, 3, or 4 as per Exemplar 5 (.95 vs .44, ($F(1, 11) = 14.40$, $MS_e = .42$ $p < .01$). This suggests that Exemplars 1, 2, 3, and 4 produced more remindings and local generalizations than did Exemplar 5 (e.g., Medin & Ross, 1989; Ross et al., 1990; Spalding &

[16] A specific comparison was only counted once regardless of how many specific features subjects mentioned (e.g., a comparison mentioning four features was only counted once). Thus, specific comparisons had the same number of opportunities as global comparisons.

Ross, 1994). As subjects became increasingly familiar with the five individuals, the number of comparisons decreased (.92 for trials 1–10, .47 for trials 11–30; $F(1, 11) = 6.63$, $MS_e = .75$, $p < .05$). This decrease was larger for specific comparisons (1.49 to .69) than for global comparisons (.35 to .25) ($F(1, 11) = 10.85$, $MS_e = .28$, $p < .01$).

The comparison data strongly suggest that Exemplar 5 served as a reference point while subjects learned the category. Because Exemplar 5 occurred so frequently, it became highly available during learning and provided a reference point for evaluating the features of less frequent exemplars. Support for this conclusion comes from the interaction in Table 7 between self versus contrast comparisons and Exemplars 1, 2, 3, and 4 versus Exemplar 5 ($F(1, 11) = 7.53$, $MS_e = .09$, $p < .05$). During trials 11 to 30, subjects performed more self comparisons for Exemplar 5 than for Exemplars 1, 2, 3, or 4 (.18 vs .04). In contrast, subjects performed more contrast trials for Exemplars 1, 2, 3, or 4 than for Exemplar 5 (.24 vs .04). Because Exemplar 5 was the object of its own self comparisons, as well as the object of contrasts to Exemplars 1, 2, 3, and 4, it was retrieved most often during learning. When Exemplar 5 was retrieved to process itself, it was strengthened in memory, enhancing its ability to function as a landmark. When Exemplar 5 was retrieved in contrast to Exemplar 1, 2, 3, or 4, it served as a standard of comparison, specifying how each of these other exemplars differed from it.

One other pattern in Table 7 further supports this conclusion. For Exemplars 1, 2, 3, and 4, there is a marginally significant trend for subjects to produce more neighbor than contrast comparisons during trials 1 to 10 (.53 vs .32) but to produce the opposite pattern during trials 11 to 30 (.18 vs .24) ($F(1, 11) = 3.59$, $MS_e = .13$, $p < .10$). During the first presentations of Exemplars 1, 2, 3, and 4, subjects were more likely to notice similarities between these individuals than to contrast them with Exemplar 5. During later presentations of Exemplars 1, 2, 3, and 4, however, subjects were more likely to contrast them with Exemplar 5 than to compare them to each other. This further suggests the increasing status of Exemplar 5 as a standard of comparison across trials.

One final analysis assessed the extent to which subjects explicitly described the stimulus structure in Table 1. Of the 12 subjects, only three noted that four individuals were similar to each other on the four shared features and differed from the frequent individual. Another three subjects partially perceived this structure, noting a subset of the relations between the four similar individuals and the frequent individual. The remaining six subjects said nothing about the stimulus structure. In a number of cases, subjects tried to extract the stimulus structure but failed, suggesting that it was difficult to see beyond the properties of a given individual and the one or two individuals brought to bear on its processing during a given trial.

Transfer protocols. During the 10 trials of the choice task, subjects justified their choices and discussed features relevant to making them (see Appendix C for examples). Overall, subjects justified a choice on 92% of the trials.

TABLE 8
Probabilities per Subject per Trial for the Content of the Transfer Protocols in Experiment 4

Probability of a choice justification	
Similar to Exemplars 1, 2, 3, and 4	.45
Similar to Exemplar 5	.18
Coherent features in the transfer exemplar	.07
Aesthetically pleasing features in the transfer exemplar	.08
Functionally useful features in the transfer exemplar	.02
Guessing	.12
Probability of mentioning a feature	
Shared dimension	.60
Unique dimension	.06
Unique dimension—irrelevant	.08

Note. The seven types of choice justifications are mutually exclusive.

Table 8 presents the probability that a subject justified a choice in one of six mutually-exclusive ways. With a probability of .63, subjects justified their choice by noting its similarity to one or more training individuals. Most often (.45), subjects justified their choice on the basis of similarity to Exemplars 1, 2, 3, and 4, stating that the 1111 transfer exemplar should be selected because it was similar to more learned individuals than was the 0000 transfer exemplar. These statements provide clear evidence that subjects often categorized on the basis of equally weighted individuals, not events. Much less often (.18), subjects based their choice on Exemplar 5, typically stating that Exemplar 5 was the most frequent individual and therefore the best basis for categorization. These statements indicate that some subjects categorized on the basis of events. However, because these subjects often referred to Exemplar 5 and the other exemplars as individuals, they based their categorizations on individuals as well. These subjects usually didn't refer just to the frequency of events—they referred to the event frequency of *individuals*. Thus, these subjects categorized on the basis of both individuals and events.

On the remaining justified trials, subjects appeared to process events independently of particular individuals. On these occasions, subjects stated that certain patterns of features seemed coherent, aesthetically pleasing, or functionally useful. Although one subject in particular generated many of these justifications, a few other subjects generated them occasionally.

Turning to the other activity that subjects performed during transfer, subjects frequently discussed features and their relevance to choices. As Table 8 illustrates, subjects mentioned a particular shared feature on a given trial with a probability of .60. Given there were four possible shared features that subjects could mention, they mentioned an average of 2.40 shared features per trial, indicating the extensive role that features played. Subjects rarely mentioned unique features, first, because all of the learned individuals differed on the unique dimensions, and second, because the two transfer exemplars on a given trial were always identical on them. Indeed, subjects were slightly more likely to state that the unique dimensions were irrelevant than that they were relevant.

Finally, subjects' protocol statements during transfer do a decent job of predicting their transfer choices. The more subjects mentioned Exemplars 1, 2, 3, and 4, and the more they discussed shared features, the more likely they were to choose 1111 transfer exemplars. The average probability across trials that a subject justified categorization on Exemplars 1, 2, 3, and 4 correlated .86 with the probability of selecting 1111 transfer exemplars minus the probability of selecting 0000 transfer exemplars. Similarly, the probability that subjects mentioned shared features correlated .88 with the 1111-0000 preference measure. We could find no strong relations between the learning protocols and transfer performance. For example, the extent to which subjects extracted the stimulus structure did not predict frequency dominance, nor did various measures of tracking, feature mention, and comparison.

Discussion

The results of Experiment 4 confirm the a priori predictions of individuals sampling models. Two variables that were predicted to affect sampling and monitoring did indeed moderate the magnitude of frequency dominance. Providing cues at transfer decreased frequency dominance, as did asking subjects to produce protocols. According to individuals sampling models, cues increased the sampling of Exemplars 1, 2, 3, and 4 through increased availability, whereas protocols increased their sampling through increased monitoring. As the sampling of Exemplars 1, 2, 3, and 4 increased, the frequency-induced dominance of Exemplar 5 decreased. Events models fail to predict these two effects. Because they don't represent individuals, they can't predict that changes in cues or monitoring should broaden the sampling of individuals during categorization.

The protocol analyses confirm and extend the conclusions reached thus far across experiments. The learning protocols clearly indicate that subjects tracked repetitions of the same individual. Rather than treating each repetition as an independent event, subjects treated them as related. The learning protocols further indicate that Exemplar 5 was not just highly available during transfer but also during learning. As subjects learned about the five individuals, they were frequently reminded of Exemplar 5, which served as a standard of comparison for encoding similarities and differences between individuals. During transfer, subjects sometimes allowed Exemplar 5's high availability to dominate categorization, justifying their choices on the basis of its high frequency. More often, subjects explicitly recognized that four similar individuals outnumbered Exemplar 5 and allowed them to control categorization.

Most importantly, subjects discussed individuals extensively throughout both learning and transfer, indicating that individuals constituted the central unit of analysis. During learning, subjects frequently recognized the same individual across repetitions as they tried to learn its features, and as they frequently compared individuals to one another. During transfer, subjects contrasted the four similar individuals with the one frequent individual. Although some subjects discussed event frequency when basing choices on Exemplar 5, most discussion proceeded with individuals as the unit of analysis.

GENERAL DISCUSSION

Prevailing Results

Across experiments, we observed frame formation, frequency dominance, appropriateness monitoring, and individual differences. Lamberts and Barsalou (1998) report further evidence for these phenomena. We discuss each in turn.

Frame formation. Subjects' frequency estimates in Experiments 1 and 2 suggested that they integrated information about the same individual across repetitions. Repeating subjects who hadn't been told the number of individuals estimated the number with reasonable accuracy. Given these estimates were far below the number of events estimated, it appears that subjects tracked individuals separately from events. Because estimates of event frequency were also reasonably accurate, subjects appeared to establish information about specific processing events as well. Interestingly, Experiment 1 observed a dissociation between individual frequency and event frequency. Subjects' estimates for the number of individuals remained constant across the equal and unequal frequency conditions, while their estimates for the number of events varied systematically with the number presented. This dissociation suggests that subjects established information about both individuals and events. In addition, a few subjects in Experiments 1 and 2 based their transfer performance on individuals, further suggesting that individuals can play important roles in categorization.

Two findings from Experiment 3 more definitively indicate that repeating subjects formed frames for individuals. First, when the frequent individual was highly dissimilar—Exemplar 5′—most subjects did not exhibit frequency dominance. For subjects to suppress Exemplar 5′, they had to realize that it was an unrepresentative individual of the category. If subjects hadn't established integrated representations of the five individuals, they wouldn't have been able to detect that Exemplar 5′ was unusual in the context of four similar individuals. In the absence of integrated representations, Exemplar 5′ should have appeared highly representative, given its 18 event memories would have seemed more representative of the category than the 12 memories of the other 4 individuals.

A second finding from Experiment 3 further implicates frame formation. Two groups of subjects received the same sequence of training exemplars that included 18 repetitions of Exemplar 5′. One group received this sequence under repeating instructions, and the other group received it under nonrepeating instructions. If repeating subjects hadn't established integrated representations of the five individuals, these subjects, like the nonrepeating subjects, should have established 18 traces for Exemplar 5′, such that it dominated categorization. However, the two groups displayed opposite preferences. The four similar individuals dominated categorization for repeating subjects, but Exemplar 5′ dominated categorization for the nonrepeating subjects. Because both groups received the same training sequence, the different instructions about individuals must have established different representations of the information in it. Whereas repeating subjects integrated the 30 stimulus events into representations for 5 individuals, nonrepeating subjects established representations for 30 individuals.

Experiment 4 offers additional evidence that subjects established frames for individuals. When subjects received cues at transfer, they were more likely to consider all five individuals during categorization. Similarly, when subjects produced protocols, they were again more likely to consider all five individuals. For both manipulations, if subjects hadn't established representations of the five individuals, they couldn't have allowed the four similar individuals to control categorization. If subjects had only stored event memories, the larger number of memories for Exemplar 5 should have dominated categorization under all conditions.

The protocols in Experiment 4 provide further evidence for frame formation. During learning and transfer, subjects tracked and compared individuals. Subjects clearly formed representations of individuals and used them as the fundamental unit of analysis.

Thus, various findings across experiments support the conclusion that subjects established frames for individuals. Theoretical considerations further suggest that integration should occur. When subjects know that individuals repeat, when they know their number, and when they know how to individuate them using names and side markings, it's difficult to imagine that they don't track them. Once an individual is recognized, previous knowledge about it should enter into processing. Top-down inferences should add information to the perceived individual, completing it and filling it in. Bottom-up input should update knowledge of the individual in memory. As knowledge of the individual interacts with the perceived individual, the two should become integrated. It is difficult to imagine how a completely independent memory for the processing episode could result from this process.

Frequency dominance. Frequency dominance occurred in at least one repeating condition of every experiment for both choice and typicality judgments. In each case, one dissimilar but highly frequent individual dominated four similar individuals. This is a surprising result, if one expects that four similar individuals should dominate one dissimilar individual. Consistent with much previous work, however, an individual's typicality in a category doesn't just depend on its featural similarity to other category members. Instead, an individual's typicality also depends on how frequently it has been encountered (e.g., Barsalou, 1981, 1985; Florian, 1992; Heit & Barsalou, 1996; Huttenlocher et al., 1996; Nosofsky, 1988; Rips & Collins, 1993).

Our interpretation of frequency dominance is that it reflects the availability of individuals during sampling (Tversky & Kahneman, 1973). The more an individual is processed during learning, the better its frame becomes established in memory. As greater amounts of knowledge become established, this information becomes increasingly integrated and strengthened, and it becomes better related to other knowledge. Together, these structural factors increase the ease and likelihood of sampling the frame, such that it dominates the frames for less established individuals. Only if additional factors make these other frames more available, such as increases in cues and monitoring, do subjects consider them.

Much remains to be learned about the availability of individuals during sampling. One issue concerns the role of automaticity (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Does a highly frequent individual dominate processing because its retrieval proceeds automatically, beyond strategic control? The fact that subjects often ignored the four similar individuals in our experiments suggests that people do not deliberate over the individuals that control their categorization decisions. Instead, people may accept the first, most dominant, individual that comes to mind effortlessly (Wright & Murphy, 1984). The fluency of this retrieval may be a contributing factor (Jacoby et al., 1989).

Another issue concerns the role of goals. When an individual is highly relevant to a current goal, does its availability increase, even if it is relatively rare? For example, the availability of poisonous mushrooms may increase in the context of picking mushrooms for a picnic repast, with automatized relations from the relevant context priming these rare individuals.

Appropriateness monitoring. As we have seen, many repeating subjects in Experiment 3 who received a frequent but highly dissimilar individual—Exemplar 5′—did not allow it to control categorization. This suggests that these subjects monitored the appropriateness of sampled individuals, recognized that Exemplar 5′ was not representative, and sampled further individuals.

When a sampled individual produces a negative monitoring response, subjects may inhibit it such that other individuals can be retrieved and contribute to the categorization decision. If such inhibition is indeed present, further issues arise. First, is this inhibition related to inhibition in other cognitive tasks (e.g., Balota & Duchek, 1991; Balota & Ferraro, 1993; Gernsbacher, 1993; Gernsbacher & Faust, 1995; Hasher & Zacks, 1988; Neely, 1977;

Posner & Snyder, 1975; Zacks & Hasher, 1994)? Second, what appropriateness criteria lead people to inhibit an individual? Experiment 3 indicates that low representativeness constitutes one criterion, as does the stereotypes literature (e.g., Kunda & Oleson, 1995, 1997). However, other possibilities exist as well, such as the contradiction of intuitive theories (Murphy & Medin, 1985). Third, what is the acceptable range of category variability within which a frequent but dissimilar individual can dominate similar but less frequent individuals? Previous research indicates that people track variability (e.g., Fried & Holyoak, 1984; Nisbett & Kunda, 1985), but more remains to be learned about the role of acceptable variability in establishing the individuals that control categorization.

Individual differences in subjects' performance. In several conditions, every subject exhibited the same preference during transfer. In other conditions, a small number of subjects differed from the majority. Consideration of these minorities suggests that subjects vary in how carefully they monitor the individuals sampled during categorization. At one extreme are subjects who monitor sampling carefully. In the repeating conditions of Experiments 1 and 2, a few subjects preferred 1111 transfer exemplars. These subjects did not need prompts, such as a highly dissimilar Exemplar 5′, cues for infrequent individuals, or instructions to produce protocols, to see the importance of exhaustive sampling. At the other extreme are subjects who do not monitor sampling at all. In the repeating / far-similarity condition of Experiment 3, for example, five subjects allowed the most available individual—Exemplar 5′—to dominate categorization, even though it was highly dissimilar. Thus, there may be a continuum of monitoring with extreme cases on each end. In the middle lie the majority of subjects, who fall under the influence of frequency in the standard repeating condition but who overcome it when a highly dissimilar individual makes the need for further sampling salient, when cues increase the availability of infrequent individuals, or when protocols increase monitoring.

A number of more specific dimensions may underlie this variability. First, subjects may vary in how carefully they update frames for individuals during learning. As each individual appears, some subjects may retrieve other individuals for comparison. If the current individual is unusual, its frame may be updated with this observation. Later, if this individual is sampled during categorization, earlier observations about its unusualness may trigger a negative monitoring response such that sampling continues. In contrast, other subjects may encode individuals less carefully, such that an unusual individual doesn't trigger a negative monitoring response. In the Experiment 4 protocols, wide variability across subjects in the production of comparisons supports this conjecture.

Second, subjects may vary in how extensively they sample individuals, either during initial sampling or after monitoring produces a negative response. Whereas some subjects may be satisfied to retrieve a single individual, other subjects may search for multiple individuals, or for significant exceptions to the individuals retrieved thus far.

Third, subjects may vary in the criteria that they consider during monitoring. Subjects may vary in their knowledge about sampling, with some subjects realizing that the more individuals retrieved, the more accurate an inference is likely to be (Smith, Langston, & Nisbett, 1992). Subjects may also vary in the category variability that they allow before further sampling is triggered. For example, fish experts may accept less variability than fish novices (or perhaps more, depending on the category).

Fourth, subjects may vary in their ability to inhibit an initially retrieved individual that is judged to be inappropriate. Failure to inhibit this individual could interfere with the further sampling of individuals, as well as with the ability to process retrieved individuals properly. Individual differences in inhibition have been found in other tasks, suggesting that such differences may also occur in the sampling of individuals during categorization (e.g., Balota & Duchek, 1991; Balota & Ferraro, 1993; Gernsbacher, 1993; Gernsbacher & Faust, 1995; Hasher & Zacks, 1988; Zacks & Hasher, 1994).

As these possibilities indicate, individuals sampling models suggest a number of individual differences that future research could pursue. During learning, subjects may vary in how carefully they update the frames for individuals. During categorization, subjects may vary in how extensively they sample individuals, in how carefully they monitor the individuals sampled, and in their ability to suppress inappropriate individuals once detected.

Theoretical Accounts

The robust frequency dominance observed across experiments clearly indicates that the individuals models in Fig. 2b are incorrect. Under many circumstances, subjects do not weight individuals equally, as these models predict. In contrast, frequency dominance supports the events models in Fig. 2a. However, evidence for the use of individuals across experiments disconfirms these models, because they lack important mechanisms for tracking and representing individuals. Instead, this complex pattern of results indicates that hybrid models are necessary. An adequate model of categorization must be capable of basing categorization on both individuals and events.

Individuals sampling models. The individuals sampling models in Fig. 5 explain the results of Experiments 1 and 2 post hoc and predict the results of Experiments 3 and 4 a priori. Furthermore, their use of availability and generate-test mechanisms is consistent with the ubiquity of these mechanisms across cognitive tasks. Thus, individuals sampling models consititute a viable account of our results.

We assume that exemplar, prototype, and connectionist theories can implement individuals sampling models. Figure 6 illustrates a local connectionist implementation, where a single net develops for each individual. In

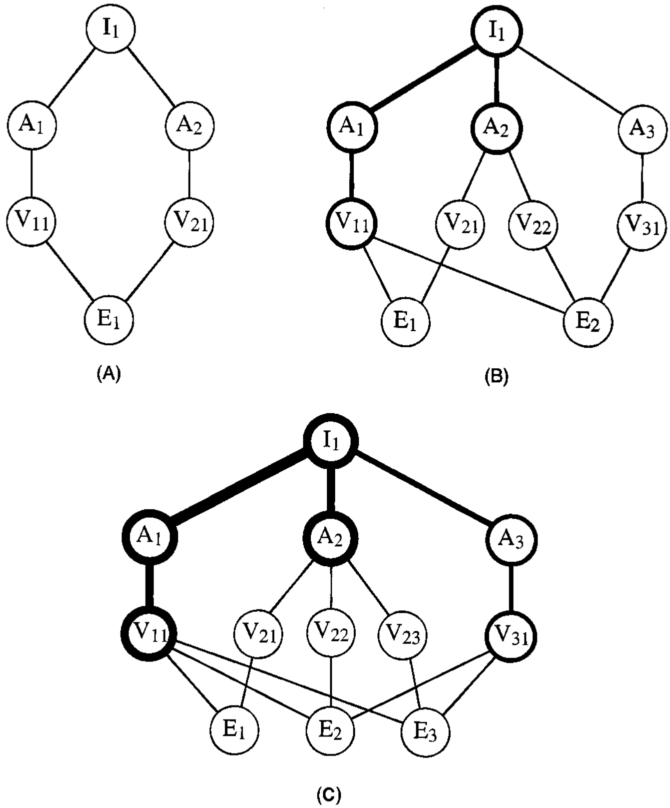


FIG. 6. The net established initially for individual I_1 after encountering it in event 1 (A), followed by the net's evolution after encountering the individual again in event 2 (B) and event 3 (C). Following each learning event, information about attributes (A_i), values (V_{ij}), and events are updated (E_i).

a first learning event (Fig. 6A), two attributes (A_1 and A_2) are encoded for individual I_1 and connected to their values (V_{11} and V_{21}), which are connected to the event (E_1). For example, the attributes might be *hair color* and *hat type*, with the values *red* and *beret*, respectively. These attributes need not be all the ones present in the individual; instead, they are the only ones noticed and encoded into memory.

In the second event (Fig. 6B), red hair color is again noticed (V_{11} for A_1), as is the presence of another hat (A_2), but this time a fedora (V_{22}). A new attribute is noticed that wasn't noticed in event 1, perhaps *eye color* (A_3) taking the value of *blue* (V_{31}). Thus, the net for I_1 is updated with two pieces of new information extracted from event 2: a new value for an old attribute, and a value for a new attribute. The strengths of repeated components and relations are also updated. As the thick circles in Fig. 6B illustrate, components repeated in event 2 are strengthened (e.g., I_1, A_1), relative to components that have only occurred once (e.g., V_{21}, A_3). As the thick lines in Fig. 6B further illustrate, connections between components repeated in event 2 are also strengthened (e.g., the connections between I_1, A_1, and V_{11}), relative to connections that have only occurred once (e.g., the connections between A_2, V_{21}, and V_{22}). As this example illustrates, the common elements of the net begin to emerge as the dominant generic structure in memory for the individual. Finally, the information in event 2 is connected to E_2, thereby representing the event. Neither this event memory, nor the one for E_1, are of the standard independent variety, however, given their integration in the net.

In the third event (Fig. 6C), red hair color is again noticed (V_{11} for A_1), as is blue eye color (V_{31} for A_3). Yet another new value is encoded for *hat*, this time a derby (V_{23} for A_2). Again, repeated components and connections are strengthened. Components and connections encoded in all three events are strongest (e.g., A_1, the relation between A_1 and V_{11}), followed by those encoded in two events (e.g., A_3, the connection between A_3 and V_{31}), followed by those encoded in one event (e.g., V_{23}, the connection between A_2 and V_{23}). Finally, the values associated with E_3 represent event 3.

This approach to representing individuals and events has several useful features. First, it produces frequency dominance. To the extent that the net for one individual accrues greater information and becomes better integrated than another, it should become more available. Second, this account supports the estimation of frequency for both individuals and events. Whereas individual frequency can be estimated by assessing the number of nets for individuals, event frequency can be estimated by assessing the number of events integrated in these nets. Third, the information processed most frequently for an individual should become most available in its net, constituting an emerging stereotype. In Fig. 6C, the stereotype is that the individual has red hair and likes to wear hats (A_1, V_{11}, and A_2). Because each type of hat was only processed once, however, *hat* does not have a stereotypical value. Finally, this account predicts that certain types of interference should occur (Anderson, 1983; Srull & Brand, 1983; Thorndyke & Hayes-Roth, 1979; Watkins & Kerkar, 1985). For example, the multiple values that fan off of A_2 should compete with each other for activation. Analogously, the multiple events that fan off of V_{11} should interfere with each other's retrieval.

Weighted individuals models. Thus far, we have assumed that the partial sampling of differentially-available individuals underlies frequency dominance. Alternatively, one could develop individuals models that sample individuals exhaustively, and that explain frequency dominance with differential weighting. Specifically, all individuals are retrieved during categorization but frequent individuals carry more weight in the decision stage than infrequent individuals. Thus, Exemplar 5 dominates Exemplars 1, 2, 3, and 4 because its individual weight is greater than their summed weights.

Although these models are feasible, they face two important challenges.

First, it seems implausible that a given person samples category information exhaustively on a given occasion, especially for well-learned categories. The literatures that we reviewed earlier on availability and generate-test mechanisms underline this concern. Availability and generate-test mechanisms wouldn't be so widespread if human memory were capable of retrieving relevant information exhaustively. The huge literature on retrieval failure further illustrates this point (e.g., Brown, 1968; Nelson, 1971, 1978; Rundus, 1973; Tulving & Pearlstone, 1966). The fact that people don't retrieve information under one set of cuing conditions (e.g., free recall) but do under another (e.g., cued recall) indicates that retrieval is far from perfect. Why should theorists should assume that retrieval is perfect in categorization when it isn't elsewhere?

Certain findings in the literature have been interpreted as indicating that subjects utilize all the information available for a category (e.g., Heit & Barsalou, 1996). Critically, however, these results rest on data pooled across subjects and trials. If different subjects sample different partial information for a category, and if the same subject samples different category information across trials, then pooling data across subjects and trials doesn't accurately portray the information that an individual subject uses in a particular judgment. Ward and Scott's (1987) analysis of holistic processing provides a sobering case in point. What appears to be holistic processing is actually an illusion that arises from pooling the data of subjects who performed analytic processing on different dimensions. When the data are pooled, it appears that subjects processed all of the dimensions holistically, when individually, they each processed one.

A second, more immediate problem for weighted individuals models is explaining the results of our experiments. How do weighted sampling models explain the distinctiveness result of Experiment 3, or the cuing and monitoring results of Experiment 4? One possibility, suggested by an anonymous reviewer, is that subjects only weight indivuals unequally when they are not distinctive. When subjects have trouble distinguishing individuals, they can't tell that they are weighting one individual more than another. In contrast, when subjects can distinguish individuals, they can tell that one individual is being weighted too heavily and then decrease its weight. Thus, subjects in Experiment 3 weighted individuals equally when Exemplar 5′ was highly distinctive but not when Exemplar 5 was undistinctive. Similarly, subjects who produced protocols in Experiment 4 were more likely to weight individuals equally, because increased monitoring led them to distinguish individuals more carefully.

Although this account explains some findings, it fails to explain others. First, in Experiment 2, as subjects received greater amounts of individuating information, frequency dominance became stronger. If increasing the distinctiveness of individuals causes them to be weighted equally, then inceasing individuating information should have weakened frequency dominance, not strengthened it. Second, in Experiment 4, subjects who received cues at transfer learned the category under identical conditions as baseline subjects who didn't receive cues. Thus, for both groups, the five individuals should have been equally distinctive after learning. However, subjects who received cues at transfer exhibited less frequency dominance. Because these cues added no new information about individuals that subjects didn't already know, distinctiveness can't explain the cuing effect. Finally, the protocols in Fig. 4 indicate that subjects mentioned the high frequency of Exemplar 5 as the basis for choosing 0000 transfer exemplars. In one example from Appendix C, a subject stated, ‘‘Well, [the 0000 transfer exemplar] looks the most like Lois [Exemplar 5], and that was the one I saw the most.’’ If subjects weight an individual equally when they can distinguish it from other individuals, then they should never state that a well-distinguished individual should be weighted the most! Clearly, however, this subject has distinguished Exemplar 5 and is explicitly weighting it more than the other individuals.

In summary, weighted individuals models face two challenges: First, they must develop an account that explains the results of these experiments. Second, they must address the larger concern about availability and generate-test mechanisms. Again, it seems quite implausible that a given person exhaustively retrieves all category knowledge on a given occasion. Instead, partial sampling seems much more likely, followed by monitoring to determine whether further sampling is necessary.

Independent trace models. Some researchers may not be comfortable with either of the models we've discussed thus far. In particular, they may be uncomfortable with integrating the repetitions of an individual into a unified representation, as in Fig. 6. Instead, these researchers may prefer approaches that keep the memory traces for a given individual separate and independent.

This approach faces several challenges. First, it must find a way of tracking and representing individuals. We are certain that such mechanisms can be added to independent trace models. We hasten to add, however, that such mechanisms are currently absent. As Figs. 2b and 5 illustrate, procedures must be developed that (a) determine whether the current individual is familiar or unfamiliar, and (b) update memory one way for familiar individuals and another way for unfamiliar individuals. For example, clusters of events could be maintained for familiar individuals, and new clusters created for unfamiliar individuals. Alternatively, learning mechanisms could insert a feature that uniquely designates each individual into its event memories, thereby distinguishing them from those for other individuals.

A second challenge for independent trace models is to develop mechanisms that categorize with events on some occasions, but with individuals on others. To account for the results of the experiments here, mechanisms at both levels are necessary. Processing event traces independently of individuals will fail under a variety of circumstances.

A final challenge for independent trace models reflects a theoretical concern. When an individual is recognized as familiar, previous knowledge should become active to process it. Perceptions are never processed in a vaccuum. As top-down inferences are made about the individual, they add features to its perceptual representation not present in bottom-up information. In turn, information extracted from perception updates the individual's previous representation. Perhaps a new feature of the individual is noticed, or perhaps some feature of the individual has changed and must be revised. As these examples illustrate, the current and previous representations of the individual participate in a mutual revision process. As a result, it seems highly likely that they become integrated, rather than ultimately residing in memory as independent traces. Clearly, individual event memories remain accessible (e.g., McCloskey & Zaragoza, 1985; Rovee-Collier, 1995). Nevertheless, it seems quite unlikely that they remain completely independent. Instead, they are likely to become integrated at one level but remain individually accessible at another, as in Fig. 6. A number of other researchers would probably agree (e.g., Adler, 1997; Medin & Ross, 1989; Ross, Perkins, & Tenpenny, 1990; Rovee-Collier, 1995; Spalding & Ross, 1994; Srull & Brand, 1983; Thorndyke & Hayes-Roth, 1979; Watkins & Kerkar, 1985).

Conclusions

In retrospect, it is not surprising that the cognitive system categorizes on the basis of both individuals and events. If the cognitive system didn't establish representations of individuals that exist across events, it couldn't construct the history of an individual, it couldn't represent the fact that the appearance of an individual might vary widely across occasions, it couldn't count the number of repeating individuals observed across occasions, and it couldn't determine the properties that occur most often across the individuals in a category. Establishing representations of individuals captures the physical structure of the world, such that important inferences about the entities in it are possible.

In contrast, if the cognitive system didn't record information about events, it couldn't distinguish individuals that occur frequently in a category from individuals that occur rarely. Similarly, it couldn't distinguish the frequent properties of an individual from the infrequent ones. In general, the representation of events captures what is likely to happen to an agent in his or her experience. Whereas frames for individuals capture what exists in the world, event memories capture how the world is likely to affect an agent in a given event.

The importance of individuals as a basic unit of cognition bears on the issue of ''representationalism'' in cognitive science. On many connectionist and dynamic systems views, distinct cognitive structures that represent individual entities in the world are unnecessary in accounting for human performance. Instead, representations that capture the statistical properties of events are sufficient. The results reported here raise problems for this approach. Without representations of individuals, an event-based system is unable to account for important regularities in human cognition. Instead, systems based on something along the lines of the one-entity one-frame assumption appear necessary (Barsalou, in press; Barsalou et. al, 1993; Johnson-Laird, 1983). Undoubtedly, a representational system must have basic statistical capabilities, such as pattern completion, generalization, frequency sensitivity, and adaptive learning. However, it must also find a way to represent entities in the world, so that reasoning can proceed at the level of individuals as well as at the level of events.

Finally, we have only addressed the possibility that the cognitive system constructs frames for individuals. However, it may also construct frames for categories. In our experiments, subjects may have established a general frame that covered all five individuals, or perhaps one that covered the four similar individuals. Barsalou et al. (1993) discuss conditions under which frames for categories may develop, such as having the explicit goal to establish generalizations that cover a set of individuals. For example, when subjects noticed that Exemplars 1, 2, 3, and 4 were similar, they may have established a frame to organize their commonalities and differences. To the extent that general frames develop, they presumably do not replace frames for individuals but instead integrate them hierarchically. The existence and structure of such frames is an open question and an important topic for future research.

APPENDIX A

Average Choice Proportions and Typicality Rankings for Individual Transfer Exemplars in Experiments 1, 2, 3, and 4

TABLE A1:
Choice Proportions and Typicality Rankings for Individual Transfer Exemplars in Experiment 1

		Unequal frequency		Equal frequency	
Exemplar	Features	Repeating	Nonrepeating	Repeating	Nonrepeating
		Choice proportions			
6	6-1111	.25	.08	.92	.92
7	6-1110	.17	.08	.92	1.00
8	6-1101	.25	.17	.67	.75
9	6-1011	.25	.08	.83	.92
10	6-0111	.17	.25	.67	.83
11	6-0000	.75	.92	.00	.00
12	6-0001	.67	.92	.08	.00
13	6-0010	.75	.92	.33	.33
14	6-0100	.75	.92	.00	.00
15	6-1000	.83	.92	.25	.17
		Typicality rankings			
6	6-1111	6.08	4.92	5.33	5.83
7	6-1110	7.25	4.58	2.83	5.75
8	6-1101	5.25	5.83	7.33	4.67
9	6-1011	7.83	8.58	4.33	3.50
10	6-0111	5.25	5.58	4.92	5.42
11	6-0000	5.33	5.92	5.33	5.17
12	6-0001	4.17	3.83	5.33	6.25
13	6-0010	4.75	7.25	6.83	3.83
14	6-0100	3.58	5.00	7.33	7.08
15	6-1000	5.50	3.33	5.42	7.50

TABLE A2:
Choice Proportions and Typicality Rankings for Individual Transfer Exemplars in Experiment 2

			Repeating/Individuation				
Exemplar	Features	Nonrepeating	None	Number	Side markings	Names	All
			Choice proportions				
6	6-1111	.00	.06	.06	.00	.31	.00
7	6-1110	.19	.25	.19	.19	.31	.13
8	6-1101	.06	.31	.50	.31	.44	.00
9	6-1011	.00	.25	.13	.00	.13	.00
10	6-0111	.00	.44	.31	.06	.38	.00
11	6-0000	1.00	.94	.94	1.00	.63	.94
12	6-0001	.88	.69	.81	.75	.63	.94
13	6-0010	.81	.75	.63	.63	.81	1.00
14	6-0100	1.00	.63	.94	1.00	.81	1.00
15	6-1000	1.00	.69	.88	1.00	.56	.94
			Typicality rankings				
6	6-1111	9.38	7.94	7.69	9.38	7.44	9.38
7	6-1110	5.81	6.94	7.19	7.63	6.06	7.69
8	6-1101	6.94	6.69	5.69	6.25	7.00	7.13
9	6-1011	8.88	7.00	7.06	8.56	7.13	8.00
10	6-0111	7.56	5.69	5.50	5.88	4.44	7.00
11	6-0000	1.63	3.00	3.75	2.19	4.13	1.50
12	6-0001	4.06	4.19	3.75	3.63	5.50	3.00
13	6-0010	4.94	4.69	5.81	5.19	5.63	4.00
14	6-0100	2.06	3.94	4.50	2.56	5.50	2.94
15	6-1000	3.75	4.88	4.06	3.75	2.19	4.25

TABLE A3:
Choice Proportions and Typicality Rankings for Individual Transfer Exemplars in Experiment 3

		Close similarity	Far similarity	
Exemplar[a]	Features	Repeating	Repeating	Nonrepeating
		Choice proportions		
6	6-1111	.00	.69	.00
7	6-1110	.06	.69	.00
8	6-1101	.06	.69	.00
9	6-1011	.00	.69	.00
10	6-0111	.00	.69	.00
11 or 11′	6-0000	1.00	.31	1.00
12 or 12′	6-0001	.88	.31	1.00
13 or 13′	6-0010	.94	.31	1.00
14 or 14′	6-0100	.94	.31	1.00
15 or 15′	6-1000	1.00	.31	1.00
		Typicality rankings		
6	6-1111	9.50	3.88	8.94
7	6-1110	7.69	4.88	8.13
8	6-1101	6.88	5.06	8.13
9	6-1011	8.50	4.13	8.19
10	6-0111	5.88	5.00	6.56
11 or 11′	6-0000	2.88	6.69	1.81
12 or 12′	6-0001	3.56	6.94	3.38
13 or 13′	6-0010	5.19	6.81	2.88
14 or 14′	6-0100	2.56	6.00	2.63
15 or 15′	6-1000	2.38	5.63	4.38

[a] As Tables 1 and 4 describe, Exemplars 11, 12, 13, 14, and 15 occurred in the close-similarity condition, whereas Exemplars 11′, 12′, 13′, 14′, and 15′ occurred in the far-similarity conditions.

TABLE A4:
Choice Proportions and Typicality Rankings for Individual Transfer Exemplars in Experiment 4

Exemplar	Features	No cues or protocols	Cues	Protocols
		Choice proportions		
6	6-1111	.00	.25	.58
7	6-1110	.00	.25	.50
8	6-1101	.06	.25	.75
9	6-1011	.00	.25	.75
10	6-0111	.00	.19	.58
11	6-0000	1.00	.75	.50
12	6-0001	1.00	.75	.42
13	6-0010	.94	.75	.33
14	6-0100	1.00	.75	.33
15	6-1000	.94	.81	.50
		Typicality rankings		
6	6-1111	9.69	6.81	4.75
7	6-1110	7.13	5.94	5.25
8	6-1101	6.38	5.31	4.83
9	6-1011	8.69	6.06	5.17
10	6-0111	6.13	6.88	4.83
11	6-0000	2.31	4.81	6.00
12	6-0001	3.00	4.69	6.58
13	6-0010	5.19	5.69	6.00
14	6-0100	2.88	4.63	5.92
15	6-1000	3.63	4.19	5.67

APPENDIX B

The Linear Exemplar Model

A linear exemplar model, along the lines suggested by Reed (1972), was developed to assess the predictions of events and individuals models. As Medin and Schaffer (1978) and Barsalou (1990) note, this linear model can be interpreted as a prototype model as well as an exemplar model. However, we interpret and implement it here using exemplars. Two versions of this model were developed. In the *weighted linear model,* each presentation of an individual establishes an independent event memory, such that frequent individuals establish more memories than infrequent individuals. In the *unweighted linear model*, only a single representation of an individual becomes established, regardless of how often it was presented. As described in Footnote 5, the weighted linear model is *not* equivalent to events models, and the unweighted linear model is *not* equivalent to individuals models. Whereas events models and individuals models are theoretical accounts of cognitive processing, the weighted and unweighted linear models are statistical tools that simply determine whether subjects weight exemplars by frequency or weight them equally.

The key measure in both models is how similar a transfer exemplar is to a stored representation of an acquisition exemplar. In the weighted linear model, the representation of an acquisition exemplar is multiple event memories established across repetitions of the same stimulus. In the unweighted linear model, there is only one representation of an acquisition exemplar, regardless of how often it was presented. In either case, the similarity of a transfer exemplar to an earlier acquisition exemplar is:

$$sim(t_i, a_j) = \sum_{d=1}^{D} w_d m_{ijd} \quad m_{ijd} = \begin{cases} 1, & t_{id} = a_{jd} \\ 0, & t_{id} \neq a_{jd} \end{cases} \tag{1}$$

The similarity of transfer exemplar t_i to acquisition exemplar a_j is the weighted sum of their dimensional matches, where D is the total number of dimensions, w_d is the weight on dimension d, and m_{ijd} is an indicator variable that is 1 when t_i and a_j match on dimension d (i.e., $t_{id} = \mathrm{a}_{jd}$), and 0 when they do not. As described in the main text, we fit three nonlinear models to the data, as well as the linear model, with the four models only differing in how they computed similarity. All other aspects of the modeling process that follow are essentially the same for the four models.

For the choice task, the overall similarity of a transfer exemplar to all of the acquisition exemplars (i.e., the category) is:

$$sim(t_i, C) = \sum_{j=1}^{A} f_{a_j}\, sim(t_i, a_j) \tag{2}$$

The similarity of transfer exemplar t_i to category C is the sum of the t_i's similarity to each of the A acquisition exemplars ($A = 5$), weighted by a_j's frequency of presentation f_{aj}, where sim(t_i, a_j) takes the form in Equation 1. For the weighted linear model, presentation frequency was 3 for Exemplars 1 through 4 versus 18 for Exemplar 5, except for the equal-frequency conditions of Experiment 1, where the presentation frequency was 3 for all exemplars (see Tables 1 and 4). For the unweighted linear model, presentation frequency was always 1 for all five exemplars.

Following Luce's choice axiom, the probability of selecting exemplar t_i of transfer pair t_i:t$_j$ is:

$$p(t_i) = \frac{sim(t_i, C)}{sim(t_i, C) + sim(t_j, C)} \tag{3}$$

Thus, the probability of choosing a member of a choice pair is the ratio of its similarity to the category to the similarities of both pair members to the category.

For the typicality task, the typicality of a transfer exemplar was its overall similarity to the category in proportion to the maximum similarity possible (i.e., a perfect match to all exemplar representations in memory):

$$typ(t_i) = \frac{sim(t_i, C)}{\max(t_*, C)} \tag{4}$$

with the maximum similarity being:

$$\max(t_*, C) = \left(\sum_{i=1}^{A} f_{a_i}\right)\left(\sum_{d=1}^{D} w_d\right) \tag{5}$$

In Equation 5, the maximum similarity of a transfer and acquisition exemplar is the sum of all possible dimension weights. Thus, the maximum similarity of a transfer exemplar to the entire category is this sum times the total number of individuals during acquisition (30 for the weighted linear model, 5 for the unweighted linear model). To invert the scale so that typicality would be correlated positively with choice probability in Equation 3, $1 - typ(t_i)$ was applied to the values obtained in Equation 4.

The dimension weights in Equation 1 were estimated directly from the data, rather than being determined by a search through the weight space for the optimal set. In fitting all four models originally, this approach ensured that all models used the same dimensional weights, rather then allowing different models to converge on different weights that range as free parameters. In the choice task, the weights are:

$$w_d = \frac{\left|\sum_{i=1}^{T} m_{id}p(t_i) - \sum_{i=1}^{T}(1 - m_{id})p(t_i)\right|}{\max(w_1, w_2, w_3, w_4)} \qquad m_{id} = \begin{cases} 1, & t_{id} = 1 \\ 0, & t_{id} = 0 \end{cases} \tag{6}$$

For dimension d, its weight, w_d, is the absolute difference of the choices for all transfer stimuli having value 1 on the dimension versus the choices for all transfer stimuli having value 0, divided by the largest dimension weight. T is the total number of transfer exemplars, and the indicator variable, m_{id}, is 1 when exemplar i has value 1 on dimension d and 0 when it has value 0. To the extent that subjects focus increasingly on a particular dimension, the absolute difference in choice between exemplars with one value versus the other should increase, assuming that subjects process the dimensions independently. Scaling the weights by the maximum absolute difference across dimensions ensures that they lie between 0 and 1.

For the typicality ratings, the weights were computed analogously to those for the choice data, except that $typ(t_i)$ replaced $p(t_i)$ in Equation 6. Dimensions shared by all transfer exemplars were not included in the fits, because they were not diagnostic.

As Tables 2, 3, 5, and 6 illustrate, the weighted and unweighted linear models produce the same absolute fit to a given data set with opposite signs (except in the equal-frequency conditions of Experiment 1, where the fits have the same absolute value and the same sign). In Table 2, for example, the unweighted model had a fit of −.84 to the repeating / unequal data, whereas the weighted model had a fit of .84. This symmetry arises because the pairwise similarities of the transfer exemplars to the training exemplars remain constant across the different implementations of frequency in the two models. To see this, note that there are two key patterns in the training exemplars: 0000, which dominates the 1111 pattern when exemplars are weighted, and 1111, which dominates the 0000 pattern when exemplars are not weighted. For the two transfer pairs containing E6 and E11, one transfer exemplar always has a 100% match with the dominant pattern, and the other has a 0% match. For the other eight transfer pairs, one transfer exemplar always has a 75% match with the dominant pattern, and the other has a 25% match. Thus, on a given transfer trial, subjects must either compare a 100% match to a 0% match, or a 75% match to a 25% match. The only difference is that these matches are defined with respect to the dominant 0000 pattern in the weighted linear model, whereas they are defined with respect to the dominant 1111 in the unweighted linear model. Because both models fit the same set of comparisons at the level of feature matches, they produce the same absolute fits. Because they assume opposite reference points—0000 versus 1111—their fits are opposite in sign. Although both models are equivalent at this level, their numerical predictions are *not* strict inverses. Because the ratio of 0000 to 1111 patterns is 18 : 12 in the nonrepeating condition but 1 : 4 in the repeating condition, the predicted transfer scores exhibit less variability in the nonrepeating condition than in the repeating condition (i.e., the 0000 and 1111 patterns are less discriminable in the former). However, this factor only causes the predictions of the two models to differ on a linear transformation. Thus, the absolute value of the correlations that assess fit are unaffected except in sign. Note that symmetry also occurs for the context, generalized context, and tuning models, when each is evaluated for a fixed set of parameter values (unlike the linear exemplar model, these models have free parameters). When parameters are free to vary between models, however, the fits are not symmetric. Different sets of parameter values optimize the weighted and unweighted models, producing different absolute fits that continue to differ in sign.

APPENDIX C

Examples of Statements That Instantiate the Content Codes in Experiment 4

In the examples to follow, Angella, Clarissa, Charmaign, and Vertna are the similar infrequent individuals, and Lois is the dissimilar frequent individual. Also, top fins and side markings are the unique features, whereas tails, bottom fins, faces, and lips are the shared features. Presentation histories and protocol histories were used to establish contextual support for codings.

Learning Protocols

Recognition. ‘‘Lois again,’’ ‘‘Lois is back again,’’ ‘‘Vertna, I’ve seen her before,’’ ‘‘Charmaign, that’s one we haven’t seen in a while,’’ ‘‘Lois is back for more.’’

Naming. ‘‘Lois,’’ ‘‘Angela,’’ ‘‘Clarissa.’’

Unique feature. ‘‘Kind of long top fin,’’ ‘‘I see Lois’ top fin, the thin fin goes up to the front,’’ ‘‘Some sort of cross-stitch pattern on the side,’’ ‘‘It has a diamond-shaped side shape.’’

Shared feature. ‘‘Spiky little tail,’’ ‘‘It’s bottom fin is hair-like,’’ ‘‘An angular face,’’ ‘‘Rosebud-shaped lips.’’

Global self-comparison. ‘‘Lois again, looks like everything’s the same,’’ ‘‘Clarissa’s back, she has the same stuff,’’ ‘‘Angela, it looks more or less the same to me.’’

Global contrast comparison. ‘‘It wasn’t the same as Lois,’’ ‘‘Contrast to Clarissa because everything is different,’’ ‘‘But it’s slightly different.’’

Global neighbor comparison. ‘‘Vertna, looks similar to the first one,’’ ‘‘Charmaign and Vertna look very similar,’’ ‘‘Looks a lot like the other ones.’’

Global previous comparison. ‘‘It looks like the same fish going through again,’’ ‘‘Contrast to Clarissa because everything’s different,’’ ‘‘Same generic design again.’’

Specific self-comparison. ‘‘I think I could recognize Lois, the same cross-stitch side pattern;’’ ‘‘It must be the same one, they’ve got the same side pattern;’’ ‘‘I recognize little lines by the attachment to the face, whatever that circular thing is.’’

Specific contrast comparison. ‘‘Something new I’ve noticed about Lois as opposed to Vertna is that Lois has a rounded face, whereas Vertna had an angular face;’’ ‘‘Round lips instead of angular lips;’’ ‘‘Lois as opposed to Angela has only one wispy thing.’’

Specific neighbor comparison. ‘‘Angela looks just like Clarissa because she has the same tail, the same bottom fin, the same face shape, the same mouth;’’ ‘‘Clarissa once again resembles Angela greatly, except that Angela has two wispy things behind her top fin and Clarissa has one in front;'' ''The attachment to the face is similar to Vertna.''

Specific previous comparison. ''Well, obviously different side pattern;'' ''Same top black thing;'' ''A different tail.''

Transfer Protocols

Similar to Exemplars 1, 2, 3, and 4. ''Four fifths of the fish had most of the characteristics present in the one on the right;'' ''Because it resembled the four fish that were not Lois, and who I consider, because they all share the same characteristics, to be more representative of the species;'' ''Because most of the fish have thin lips, a single hair for a bottom fin, and this kind of beak-shaped tail fin;'' ''It has the other three sort of normal characteristics, or what seemed normal for the fishes that weren't Lois;'' ''Well Lois was the only fish that looked like the fish on the left. Most of the fish in the species seemed to look like the fish on the right.''

Similar to Exemplar 5. ''The one on the right looks most like Lois again, because that's the one that I'm remembering;'' ''Well, that looks the most like Lois, and that was the one I saw the most;'' ''I'd have to say that the one on the left seems more likely because of the fact that I saw Lois so many times;'' ''I'd say the second one because it looks a lot like Lois;'' ''Because I've seen those shapes more often than the left side.''

Coherent features in the transfer exemplar. ''The cone-shaped head and the thin lips go with the head that has a crown to the right of the eyes;'' ''The single fins on the bottom tend to go with the really small mouth.''

Aesthetically pleasing features in the transfer exemplar. ''I do prefer the single line to the multiple lines,'' ''I really don't like those lips.''

Functionally useful features in the transfer exemplar. ''I like the rescue ranger aspect of the fish that we saw, the little thing on the bottom, and I think that that's an important part of their psychological environment;'' ''I think that's more important for their psychological habitat.''

Guessing. ''I really don't know why I'm making this decision,'' ''At this point I feel like I'm guessing instead of remembering any special features of certain fish.''

Shared feature. ''I'll probably have to pick the one on the left because it has a single bottom fin, the sea-shaped instead of jagged tail fin, the beak mouth;'' ''The one with the mouth like four of the fish;'' ''And only Lois had the big mouth and the small back fin;'' ''Because of the bottom fin.''

Unique feature. ''I choose this one to be more likely because the top fin doesn't have any breaks in it;'' ''I'd say the second one because other than Lois the other fish had different top fins than this.''

Unique feature—irrelevant. ''Side patterns are different for each fish, so that doesn't matter;'' ''We don't seem to care about the top fin.''

REFERENCES

Adler, S. A. (1997). Selective integration of temporally distinct perceptual information in infant memory. *Infant Behavior and Development,* **20,** 3–28.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review,* **79,** 97–123.

Andersen, S. M., & Cole, S. W. (1990). ''Do I know you?'': The role of significant others in general social perception. *Journal of Personality and Social Psychology,* **59,** 384–399.

Andersen, S. M., & Glassman, N. S. (1996). Responding to significant others when they are not there: Effects on interpersonal inference, motivation, and affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 3, pp. 272–331). New York: Guilford Press.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Armstrong, D. M. (1989). *Universals: An opinionated introduction.* Boulder, CO: Westview Press.

Balota, D. A., & Duchek, J. M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. *Brain and Language,* **40,** 181–201.

Balota, D. A., & Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language,* **32,** 573–592.

Bareiss, R., & Slator, B. M. (1993). The evolution of a case-based computational approach to knowledge, representation, classification, and learning. In G. V. Nakamura, D. L. Medin, & R. Taraban (Eds.), *Categorization by humans and machines. Special volume of The Psychology of Learning and Motivation* (Vol. 29, 157–186). San Diego, Academic Press.

Barsalou, L. W. (1981). *Determinants of graded structure in categories.* Unpublished doctoral dissertation. Stanford University.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **11,** 629–654.

Barsalou, L. W. (1988). The content and organization of autobiographical memories. In U. Neisser (Ed.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 193–243). New York: Cambridge University Press.

Barsalou, L. W. (1992). Frames, concepts, and conceptual fields. In E. Kittay & A. Lehrer (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 21–74). Hillsdale, NJ: Erlbaum.

Barsalou, L. W. (1993). Structure, flexibility, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. C. Collins, S. E. Gathercole, & M. A. Conway (Eds.), *Theories of memory* (pp. 29–101). London: Lawrence Erlbaum Associates.

Barsalou, L. W. (in press). Perceptual symbol systems. *Behavioral and Brain Sciences.*

Barsalou, L. W., Yeh, W., Luka, B. J., Olseth, K. L., Mix, K. S., & Wu, L. (1993). Concepts and meaning. In In K. Beals, G. Cooke, D. Kathman, K. E. McCullough, S. Kita, & D. Testen (Eds.), *Chicago Linguistics Society 29: Papers from the parasession on conceptual representations* (pp. 23–61). University of Chicago: Chicago Linguistics Society.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* New York: Cambridge University Press.

Bauer, P. J., & Fivush, R. (1992). Constructing event representations: Building on a foundation of variation and enabling relations. *Cognitive Development,* **7,** 381–401.

Beike, D. R., & Sherman, S. J. (1994). Social inference: Inductions, deductions, and analogies. In R. S. Wyer, Jr. & Srull, T. K. (Eds.), *Handbook of social cognition: Vol. 1. Basic processes* (2nd ed., pp. 209–285). Mahwah, NJ: Lawrence Erlbaum Associates.

Bloom, P., & Kelemen, D. (1995). Syntactic cues in the acquisition of collective nouns. *Cognition,* **56,** 1–30.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (169–211). Hillsdale, NJ: Erlbaum.

Brown, J. (1968). Reciprocal facilitation and impairment of free recall. *Psychonomic Science,* **10,** 41–42.

Bruner, J. S., & Potter, M. C. (1964). Interference in visual recognition. *Science,* **144,** 424–425.

Ceraso, J., & Provitera, A. (1971). Sources of error in syllogistic reasoning. *Cognitive Psychology,* **2,** 400–410.

Estes, W. K. (1959). The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 380–491). New York: McGraw-Hill.

Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development,* **63,** 173–187.

Florian, J. (1992). Frequency effects on categorization and recognition. *Proceedings of the 14th annual conference of the Cognitive Science Society,* 826–831. Hillsdale, NJ: Erlbaum.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **10,** 234–257.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar.* New York: McGraw-Hill.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science,* **4,** 294–298.

Gernsbacher, M. A., & Faust, M. (1995). Skilled suppression. In F. N. Dempster & C. N. Brainerd (Eds.), *New perspectives on interference and inhibition in cognition* (pp. 295–327). New York: Academic Press.

Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review,* **102,** 4–27.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and new view. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol 22* (pp. 193–223). New York: Academic Press.

Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 1264–1282.

Heit, E., & Barsalou, L. W. (1996). The instantiation principle in natural categories. *Memory,* **4,** 413–451.

Hintzman, D. L. (1986). ''Schema abstraction'' in a multiple-trace memory model. *Psychological Review,* **93,** 411–428.

Huttenlocher, J., Hedges, L., Engebretson, H., & Vevea, J. (1996). *Reconstructing experiences from memory: The effects of categories.* Manuscript under review.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology,* **56,** 326–338.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior,* **23,** 734–759.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review,* **99,** 22–44.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of discrimination: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology,* **68,** 565–579.

Kunda, Z., & Oleson, K. (1997). When exceptions prove the rule: How extremity of deviance determines deviants' impact on stereotypes. *Journal of Personality and Social Psychology,* **72,** 965–979.

Lamberts, K. (1994). Flexible tuning of similarity in exemplar-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 1003–1021.

Lamberts, K., & Barsalou, L. W. (1998). *Exemplar sampling and frequency effects in category learning.* In preparation.

Luchins, A. S. (1942). Mechanization in problem solving. *Psychological Monographs,* **54** (Whole No. 248).

McClelland, J. L., Rumelhart, D. E., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models.* Cambridge, MA: MIT Press.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General,* **114,** 1–16.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General,* **117,** 68–85.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem-solving, and induction. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5 pp. 189–223). Hillsdale, NJ: Erlbaums.

Medin, D. L., & Schaffer, M. (1978). A context theory of classification learning. *Psychological Review,* **85,** 207–238.

Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs, and real kinds: More Mama, more milk, more mouse. *Behavioral and Brain Sciences,* **21,** 55–100.

Minksy, M. L. (1977). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York: McGraw-Hill.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review,* **92,** 289–316.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory. Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General,* **106,** 226–254.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nelson, T. O. (1971). Savings and forgetting from long-term memory. *Journal of Verbal Learning and Verbal Behavior,* **10,** 568–576.

Nelson, T. O. (1978). Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **4,** 453–368.

Nisbett, R. E., & Kunda, Z. (1985). Perception of social distributions. *Journal of Personality and Social Psychology,* **48,** 297–311.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. A., Rumelhart, D. E., & the LNR Research Group. (1975). *Explorations in cognition.* New York: Freeman.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **10,** 104–114.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14,** 54–65.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance,* **17,** 3–27.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review,* **104,** 266–300.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology,* **3,** 382–407.

Rips, L. J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology: General,* **122,** 468–486.

Ross, B. H., Perkins, S. J., & Tenpenny, P. L. (1990). Reminding-based category learning. *Cognitive Psychology,* **22,** 460–492.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, P. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology,* **14,** 237–255.

Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology,* **31,** 147–169.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol 1. Foundations.* Cambridge, MA: MIT Press.

Rumelhart, D. E., & Ortony, A. (1978). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior,* **12,** 43–50.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review,* **84,** 1–66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review,* **84,** 127–190.

Smith, E. E., Langston, C., & Nisbett, R. (1992). The case for rules in reasoning. *Cognitive Science,* **16,** 1–40.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. E., & Sloman, S. A. (1994). Similarity- vs. rule-based categorization. *Memory & Cognition,* **22,** 377–386.

Spalding, T. L., & Ross, B. H. (1994). Comparison-based learning: Effects of comparing instances during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 1251–1263.

Srull, T. K., & Brand, J. F. (1983). Memory for information about persons: The effect of encoding operations on subsequent retrieval. *Journal of Verbal Learning and Verbal Behavior,* **22,** 219–230.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review,* **96,** 58–83.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior,* **18,** 645–659.

Thorndyke, P. W., & Hayes-Roth, B. (1979). The use of schemata in the acquisition and transfer of knowledge. *Cognitive Psychology,* **11,** 82–106.

Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychologist,* **47,** 862–875.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior,* **5,** 381–391.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* **5,** 207–232.

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family resemblance concepts. *Memory & Cognition,* **15,** 42–54.

Watkins, M. J., & Kerkar, S. P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language,* **24,** 666–678.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill Book Company.

Wright, J. C., & Murphy, G. L. (1985). The utility of theories in intuitive statistics: The robustness of theory-based judgments. *Journal of Experimental Psychology: General,* **113,** 301–322.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology,* **30,** 111–153.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory mechanisms in attention, memory, and language* (pp. 241–264). New York: Academic Press.

Accepted May 7, 1998